MODERN SPECTRUM ANALYSIS

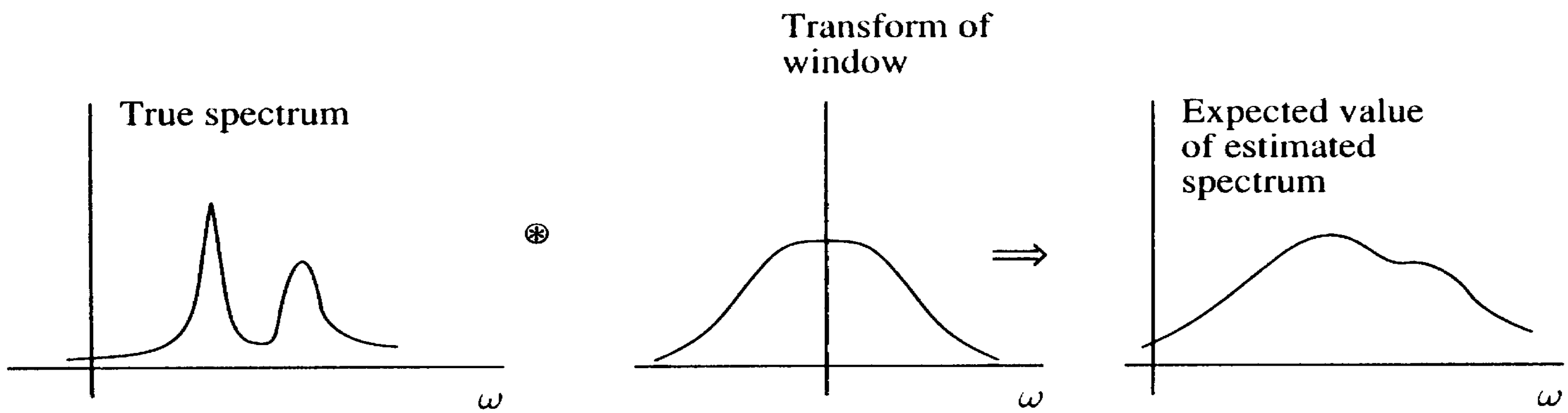
- Methods based on linear models
  - AR
  - MA
  - ARMA
- AR method and Maximum Entropy
- “Maximum Likelihood” method

MODERN SPECTRUM ANALYSIS (cont'd.)

- Subspace methods
  - Pisarenko
  - MUSIC
  - Minimum Norm
  - Principal Components Linear Prediction
  - *ESPRIT*

LIMITATIONS OF CLASSICAL METHODS

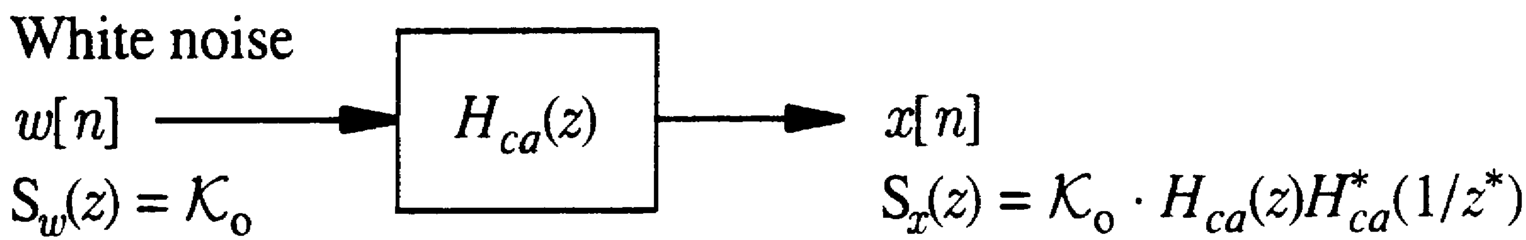
- Classical methods are limited in resolution by the data length.



- Methods based on a model for the process can overcome this limitation.

SPECTRAL ESTIMATION USING A LINEAR MODEL

MODEL FOR THE PROCESS



SPECTRAL ESTIMATE

$$\widehat{S}(e^{\jmath\omega}) = \mathcal{K}_o \, |H_{ca}(e^{\jmath\omega})|^2$$

FORMS OF SPECTRAL ESTIMATES

AR

$$\widehat{S}_{AR}(e^{\jmath\omega}) = \frac{|b_0|^2}{|A(e^{\jmath\omega})|^2} = \frac{\sigma_P^2}{|A(e^{\jmath\omega})|^2}$$

MA

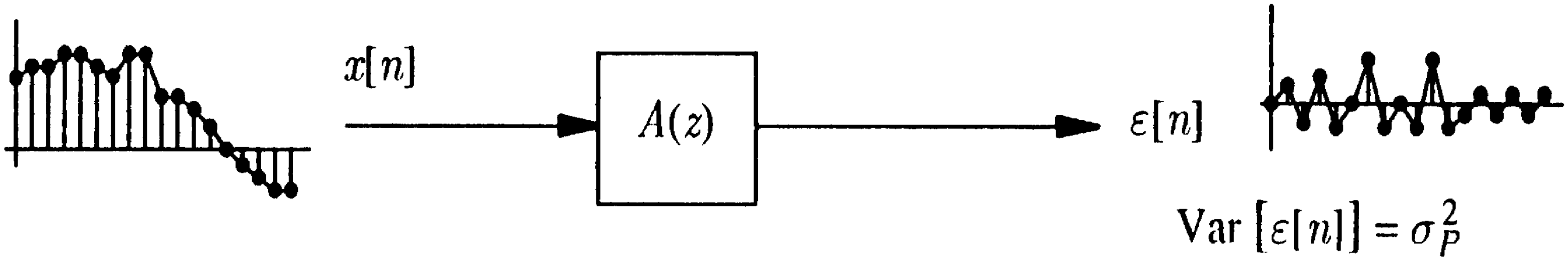
$$\widehat{S}_{MA}(e^{\jmath\omega}) = |B(e^{\jmath\omega})|^2$$

ARMA

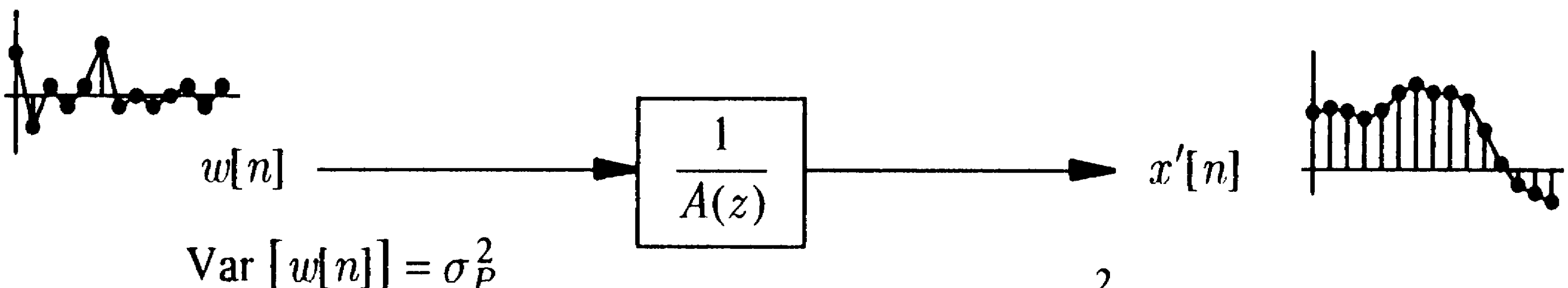
$$\widehat{S}_{ARMA}(e^{\jmath\omega}) = \left|\frac{B(e^{\jmath\omega})}{A(e^{\jmath\omega})}\right|^2$$

SPECTRUM ESTIMATION BY AR MODELING

LINEAR PREDICTION



AR MODELING



$$S_x(e^{j\omega}) = S_{x'}(e^{j\omega}) = \frac{\sigma_P^2}{|A(e^{j\omega})|^2}$$

PROPERTIES OF THE AR MODEL

CORRELATION MATCHING

$$R_{x'}[l] = R_x[l] \; ; \qquad l = 0, \pm 1, \pm 2, \ldots, \pm P$$

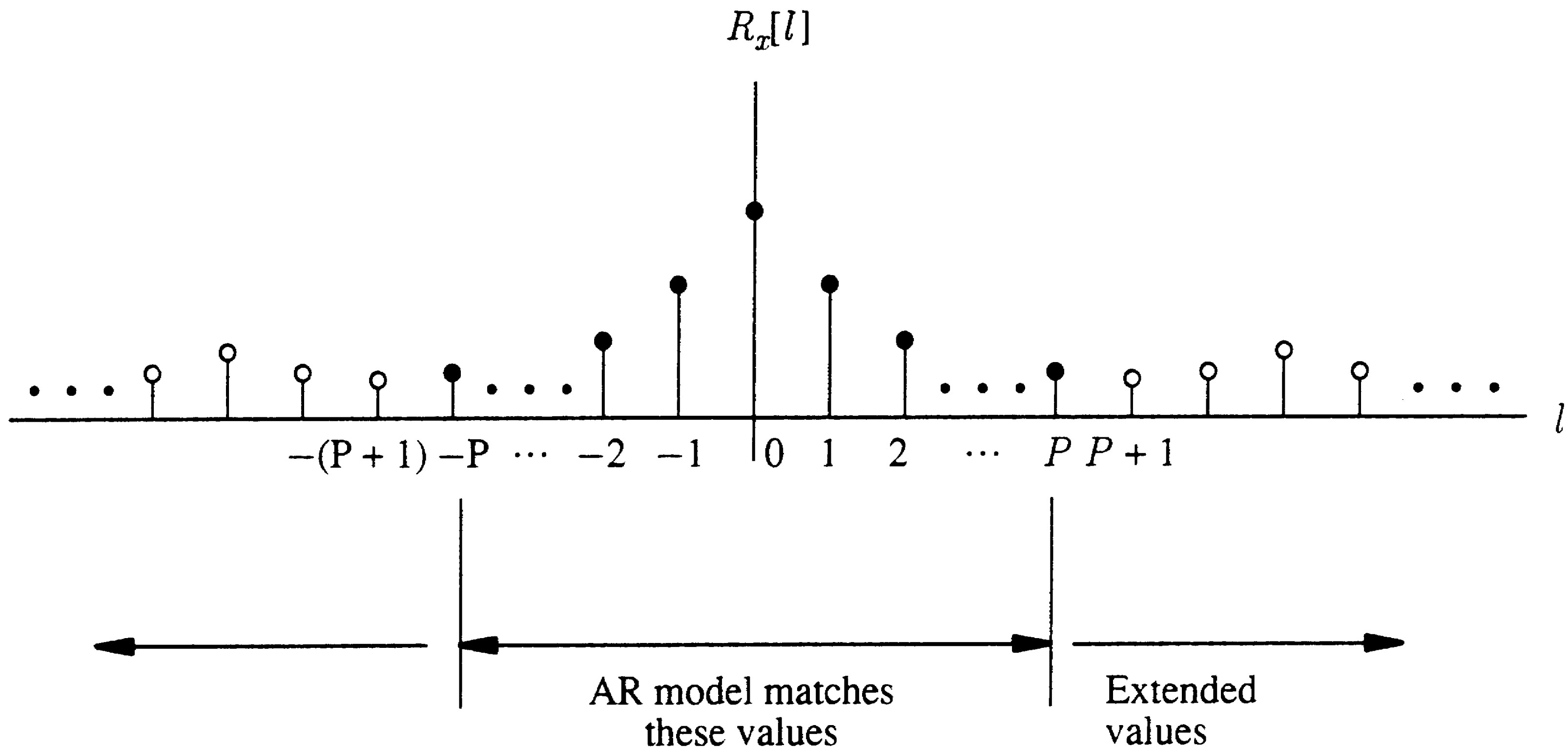
CORRELATION EXTENSION

$$R_{x'}[l] + a_1 R_{x'}[l-1] + a_2 R_{x'}[l-2] + \cdots + a_P R_{x'}[l-P] = \underbrace{R_{wx'}[l]}_{0 \text{ for } l>0}$$

$\Longrightarrow$

$$R_{x'}[l] = -a_1 R_{x'}[l-1] - a_2 R_{x'}[l-2] - \cdots - a_P R_{x'}[l-P], \qquad l > 0$$

MATCHING AND EXTENSION OF THE CORRELATION FUNCTION



MAXIMUM ENTROPY PROPERTY

- The AR model of order P matches the correlation function up to lag P.

- The AR model extends the correlation function in a way to maximize entropy of the resulting process.

In other words . . .

- Of all processes that could match and extend the given correlation function, the AR process is the process with *maximum entropy*.

PROOF OF MAXIMUM ENTROPY

The entropy for $p+1$ samples of a zero-mean complex Gaussian random process is

$$\mathcal{H}_p = \mathcal{E}\left\{-\ln f_{\boldsymbol{x}_p}(\boldsymbol{x}_p)\right\} = (p+1)(1+\ln\pi) + \ln|\mathbf{R}_{\boldsymbol{x}}^{(p)}|$$

where

$$\mathbf{R}_{\boldsymbol{x}}^{(p)} = \begin{bmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-p] \\ R_x[1] & R_x[0] & \cdots & R_x[-p+1] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[p] & R_x[p-1] & \cdots & R_x[0] \end{bmatrix}$$

- Have terms $R_x[0], R_x[1], \ldots, R_x[P]$.

- Need to choose $R_x[P+1], R_x[P+2], \ldots$ to maximize $\mathcal{H}_p$ for all values of p.

PROOF OF MAXIMUM ENTROPY (cont'd.)

First choose $R_x[P+1]$ to maximize $|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}|$:

$$
\begin{aligned}
|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}| &= \begin{vmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-P] & R_x[-P-1] \\ R_x[1] & R_x[0] & \cdots & R_x[-P+1] & R_x[-P] \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ R_x[P] & R_x[P-1] & \cdots & R_x[0] & R_x[-1] \\ R_x[P+1] & R_x[P] & \cdots & R_x[1] & R_x[0] \end{vmatrix} \\
&= R_x^*[P+1](-1)^{P+1} \begin{vmatrix} R_x[1] & R_x[0] & \cdots & R_x[-P+1] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P] & R_x[P-1] & \cdots & R_x[0] \\ R_x[P+1] & R_x[P] & \cdots & R_x[1] \end{vmatrix} \\
&\qquad + \text{other terms}
\end{aligned}
$$

PROOF OF MAXIMUM ENTROPY (cont'd.)

A necessary condition for the maximum is

$$\nabla_{R_x^*[P+1]}|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}| = (-1)^{P+1} \begin{vmatrix} R_x[1] & R_x[0] & \cdots & R_x[-P+1] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P] & R_x[P-1] & \cdots & R_x[0] \\ R_x[P+1] & R_x[P] & \cdots & R_x[1] \end{vmatrix} = 0$$

It can be shown that this condition *indeed* produces a maximum, i.e.,

$$\left(\nabla^2_{R_x R_x}|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}|\right) \cdot \left(\nabla^2_{R_x^* R_x^*}|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}|\right) - \left(\nabla^2_{R_x R_x^*}|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}|\right)^2 < 0$$

and

$$\nabla^2_{R_x R_x^*}|\mathbf{R}_{\boldsymbol{x}}^{(P+1)}| < 0$$

(see text).

PROOF OF MAXIMUM ENTROPY (cont'd.)

Since the determinant is zero, the columns are linearly dependent.

$$\begin{bmatrix} R_x[1] & R_x[0] & \cdots & R_x[-P+1] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P] & R_x[P-1] & \cdots & R_x[0] \\ R_x[P+1] & R_x[P] & \cdots & R_x[1] \end{bmatrix} \begin{bmatrix} 1 \\ c_1 \\ \vdots \\ c_P \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

If a top row is added...

$$\begin{bmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-P] \\ R_x[1] & R_x[0] & \cdots & R_x[-P+1] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P] & R_x[P-1] & \cdots & R_x[0] \\ ----- & ------ & -- & ------- \\ R_x[P+1] & R_x[P] & \cdots & R_x[1] \end{bmatrix} \begin{bmatrix} 1 \\ c_1 \\ \vdots \\ c_P \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ \vdots \\ 0 \\ - \\ 0 \end{bmatrix}$$

this defines the AR model and extension of R_x to $R_x[P+1]$.

PROOF OF MAXIMUM ENTROPY (cont'd.)

A repetition of the analysis for extension to $R_x[P+2]$ leads to

$$\begin{bmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-P-1] \\ R_x[1] & R_x[0] & \cdots & R_x[-P] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P+1] & R_x[P] & \cdots & R_x[0] \\ ----- & ----- & -- & ----- \\ R_x[P+2] & R_x[P+1] & \cdots & R_x[1] \end{bmatrix} \begin{bmatrix} 1 \\ c'_1 \\ \vdots \\ c'_P \\ c'_{P+1} \end{bmatrix} = \begin{bmatrix} \sigma'^2 \\ 0 \\ \vdots \\ 0 \\ -- \\ 0 \end{bmatrix}$$

This can be satisfied by taking ${c'}_1 = c_1, \ldots, {c'}_P = c_P, {c'}_{P+1} = 0$. The bottom row then provides the correlation extension

$$R_x[P+2] = -c_1 R_x[P+1] - \cdots - c_P R_x[2]$$

This procedure is continued to find $R_x[P+3]$, $R_x[P+4]$, ...

BURG'S PROOF OF MAXIMUM ENTROPY

- Start with the entropy (per sample) for a Gaussian process

$$\Delta\mathcal{H} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln S_{x'}(e^{\jmath\omega})d\omega \ + \text{const.}$$

- Maximize this subject to constraints

$$R_{x'}[l] = \frac{1}{2\pi}\int_{-\pi}^{\pi} S_{x'}(e^{\jmath\omega})e^{\jmath\omega l}d\omega = R_x[l] \ ; \quad l = 0, \pm 1, \ldots, \pm P$$

- Show that an all-pole model is required

- Show that the all-pole model is the AR model

BURG'S PROOF (cont'd.)

The necessary condition for a maximum is

$$\nabla_{R^*_{x'}[l]} \Delta\mathcal{H} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{S_{x'}(e^{\jmath\omega})} \left(\nabla_{R^*_{x'}[l]} S_{x'}(e^{\jmath\omega}) \right) d\omega = 0 \ ; \quad |l| > P$$

Note that $S_{x'}(e^{\jmath\omega})$ can be written as

$$\sum_{k=-\infty}^{\infty} R_{x'}[k] e^{-j\omega k} = \sum_{k'=-\infty}^{\infty} R^*_{x'}[k'] e^{j\omega k'} \quad \Longrightarrow \quad \nabla_{R^*_{x'}[l]} S_{x'}(e^{\jmath\omega}) = e^{\jmath\omega l}$$

Therefore

$$\nabla_{R^*_{x'}[l]} \Delta\mathcal{H} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{S_{x'}(e^{\jmath\omega})} e^{\jmath\omega l} d\omega = 0 \ ; \quad |l| > P$$

BURG'S PROOF (cont'd.)

The condition

$$\nabla_{R^*_{x'}[l]} \Delta\mathcal{H} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{S_{x'}(e^{\jmath\omega})} e^{\jmath\omega l} d\omega = g[l] = 0 \ ; \qquad |l| > P$$

states that the sequence $g[l]$ with transform $1/S_{x'}(z)$ has finite length, *i.e.*,

$$S_{x'}(z) = \frac{1}{\sum_{l=-P}^{P} g[l] z^{-l}} = \frac{\mathcal{K}_0}{\left(\sum_{n=0}^{P} c_n z^{-n}\right)\left(\sum_{k=0}^{P} c_k^* z^k\right)} \qquad (c_0 = 1)$$

Therefore the process can be modeled as white noise driving the all-pole filter

$$H_{ca}(z) = \frac{1}{\sum_{n=0}^{P} c_n z^{-n}}$$

BURG'S PROOF (cont'd.)

The form of the power spectral density implies

$$
\begin{aligned}
\frac{\mathcal{K}_o}{\sum_{n=0}^{P} c_n z^{-n}} &= \left(\sum_{k=0}^{P} c_k^* z^k\right) S_{x'}(z) = \sum_{k=0}^{P} c_k^* z^k \sum_{l=-\infty}^{\infty} R_{x'}[l] z^{-l} \\
&= \sum_{l'=-\infty}^{\infty} \left(\sum_{k=0}^{P} c_k^* R_{x'}[l'+k]\right) z^{-l'} \qquad\qquad (l' = l-k) \\
&= \sum_{l'=-\infty}^{\infty} \underbrace{\left(\sum_{k=0}^{P} c_k R_{x'}[-l'-k]\right)^*}_{\substack{0 \text{ for } l' < 0 \\ \mathcal{K}_o \text{ for } l' = 0}} z^{-l'}
\end{aligned}
$$

BURG'S PROOF (cont'd.)

Finally, the last condition

$$\sum_{k=0}^{P} c_k R_{x'}[-l'-k] = \begin{cases} \mathcal{K}_{\mathrm{o}} & \text{for } l'=0 \\ 0 & \text{for } l'<0 \end{cases}$$

and the requirement $R_{x'}[l] = R_x[l]$ for $|l| \leq P$ produces the Yule-Walker equations for the AR model

$$\begin{bmatrix} R_x[0] & R_x[-1] & \cdots & R_x[-P] \\ R_x[1] & R_x[0] & \cdots & R_x[-P+1] \\ \vdots & \vdots & \vdots & \vdots \\ R_x[P] & R_x[P-1] & \cdots & R_x[0] \end{bmatrix} \begin{bmatrix} 1 \\ c_1 \\ \vdots \\ c_P \end{bmatrix} = \begin{bmatrix} \mathcal{K}_{\mathrm{o}} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

“MAXIMUM ENTROPY” SPECTRUM ESTIMATION

- “Maximum Entropy” is the name given by Burg to his method of spectrum estimation.

- Theoretically, any AR spectral estimate is a maximum entropy spectral estimate.

- In practice, the term is reserved to mean an AR estimate where the model parameters are computed using Burg’s method.

COMPUTATION OF SPECTRAL ESTIMATES

- Efficient model-based spectral estimates can be computed with the FFT.

- Note that

$$\widehat{S}(e^{\jmath\omega}) = \left|\frac{B(e^{\jmath\omega})}{A(e^{\jmath\omega})}\right|^2 = \frac{|\text{FT of sequence } \{b_n\}|^2}{|\text{FT of sequence } \{a_n\}|^2}$$

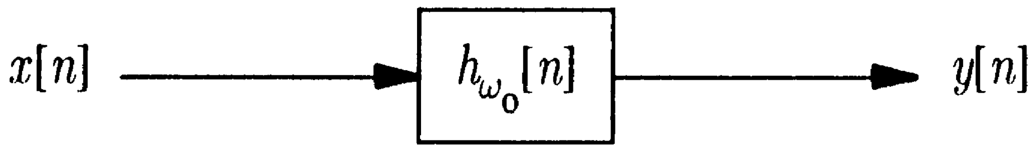
- The FT can be replaced by the DFT computed with an FFT program; for purposes of ploting, the DFT order is chosen to give a smooth plot.

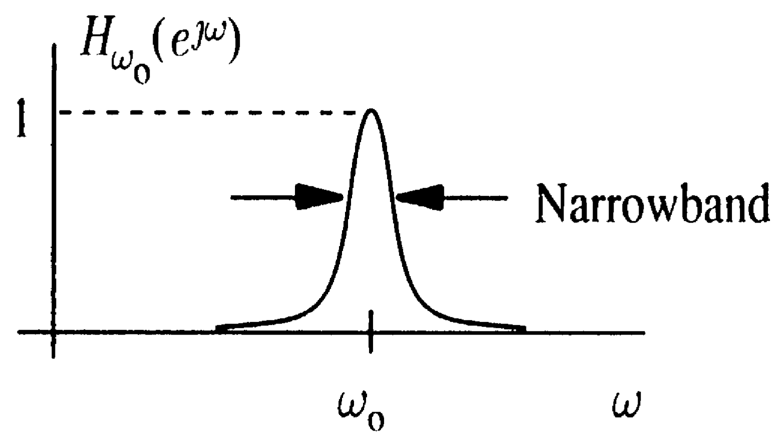
“MAXIMUM LIKELIHOOD” SPECTRUM ESTIMATION

- Computed using the correlation matrix for the data
- Exhibits higher resolution than classical methods (Bartlett, Blackman-Tukey, etc.)
- Can be related to the “Maximum Entropy” method
- Also called the “Minimum Variance (Distortionless)” method

"MAXIMUM LIKELIHOOD" METHOD

SPECTRUM ANALYZER INTERPRETATION





ML spectral estimate: $\widehat{S}_{ML}(e^{j\omega_o}) \stackrel{\text{def}}{=} \mathcal{E}\left\{|y[n]|^2\right\}$

ML METHOD DEVELOPMENT

The FIR filter output is

$$y[n] = \sum_{k=0}^{N-1} h_{\omega_0}[k]x[n-k] = \mathbf{h}_0^T \tilde{\boldsymbol{x}}[n]$$

The filter average power is given by

$$\mathcal{P} = \mathcal{E}\left\{|y[n]|^2\right\} = \mathbf{h}_0^T \mathcal{E}\left\{\tilde{\boldsymbol{x}}[n]\tilde{\boldsymbol{x}}^{*T}[n]\right\}\mathbf{h}_0^* = \mathbf{h}_0^T \tilde{\mathbf{R}}_{\boldsymbol{x}} \mathbf{h}_0^* = \mathbf{h}_0^{*T} \mathbf{R}_{\boldsymbol{x}} \mathbf{h}_0$$

This is minimized subject to the (complex) constraint

$$H_{\omega_0}(e^{\jmath\omega_0}) = \sum_{n=0}^{N-1} h_{\omega_0}[n] e^{-\jmath\omega_0 n} = \mathbf{w}_0^{*T} \mathbf{h}_0 = 1$$

ML METHOD (cont'd.)

The optimization problem involves the Lagrangian

$$\mathcal{L} = \mathbf{h}_\mathrm{O}^{*T}\mathbf{R}_{\boldsymbol{x}}\mathbf{h}_\mathrm{O} + \mu(1 - \mathbf{w}_\mathrm{O}^{*T}\mathbf{h}_\mathrm{O}) + \mu^*(1 - \mathbf{h}_\mathrm{O}^{*T}\mathbf{w}_\mathrm{O})$$

and the necessary condition

$$\nabla_{\mathbf{h}_\mathrm{O}^*}\mathcal{L} = \mathbf{R}_{\boldsymbol{x}}\mathbf{h}_\mathrm{O} - \mu^*\mathbf{w}_\mathrm{O} = \mathbf{0} \qquad \Longrightarrow \qquad \mathbf{h}_\mathrm{O} = \mu^*\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_\mathrm{O}$$

The requirement $\mathbf{w}_\mathrm{O}^{*T}\mathbf{h}_\mathrm{O} = \mu^*\mathbf{w}_\mathrm{O}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_\mathrm{O} = 1$ then yields

$$\mu^* = \mu = \frac{1}{\mathbf{w}_\mathrm{O}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_\mathrm{O}} \quad \text{so that} \quad \mathbf{h}_\mathrm{O} = \frac{\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_\mathrm{O}}{\mathbf{w}_\mathrm{O}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_\mathrm{O}}$$

ML METHOD (cont'd.)

The optimum narrowband filter at frequency ω_o

$$\mathbf{h}_o = \frac{\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_o}{\mathbf{w}_o^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_o}$$

produces the output power

$$\mathcal{P} = \mathbf{h}_o^{*T}\mathbf{R}_{\boldsymbol{x}}\mathbf{h}_o = \frac{\mathbf{w}_o^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{R}_{\boldsymbol{x}}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_o}{(\mathbf{w}_o^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_o)^2} = \frac{1}{\mathbf{w}_o^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}_o}$$

This is the ML power spectral estimate of the process at frequency ω_o.

“MAXIMUM LIKELIHOOD” METHOD (SUMMARY)

The “Maximum Likelihood” spectral estimate is

$$\boxed{\widehat{S}_{ML}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}}}$$

where

$$\mathbf{w} = \begin{bmatrix} 1 \\ e^{\jmath\omega} \\ e^{\jmath 2\omega} \\ \vdots \\ e^{\jmath(N-1)\omega} \end{bmatrix}$$

CLASSICAL METHOD COMPARED TO "MAXIMUM LIKELIHOOD" METHOD

If the Fourier transform of the data sequence is

$$X(e^{\jmath\omega}) = \sum_{n=0}^{N-1} x[n]e^{-\jmath\omega n} = \mathbf{w}^{*T}\boldsymbol{x}$$

the periodogram spectral estimate is defined by

$$\widehat{P}_x(e^{\jmath\omega}) \overset{\text{def}}{=} \frac{1}{N}|X(e^{\jmath\omega})|^2 = \frac{1}{N}X(e^{\jmath\omega})X^*(e^{\jmath\omega}) = \frac{1}{N}\mathbf{w}^{*T}\boldsymbol{x}\boldsymbol{x}^{*T}\mathbf{w}$$

The expected value of this estimate is

$$\mathcal{E}\left\{\widehat{P}_x(e^{\jmath\omega})\right\} = \frac{1}{N}\mathbf{w}^{*T}\mathbf{R}_{\boldsymbol{x}}\mathbf{w} \quad \text{while} \quad \widehat{S}_{ML}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}}$$

RELATION BETWEEN ML AND ME

Use the triangular decomposition to write

$$\widehat{S}_{ML}^{-1}(e^{\jmath\omega}) = \mathbf{w}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w} = \mathbf{w}^{*T}(\mathbf{U}_1^{-1})^{*T}\mathbf{D}_U^{-1}\mathbf{U}_1^{-1}\mathbf{w}$$

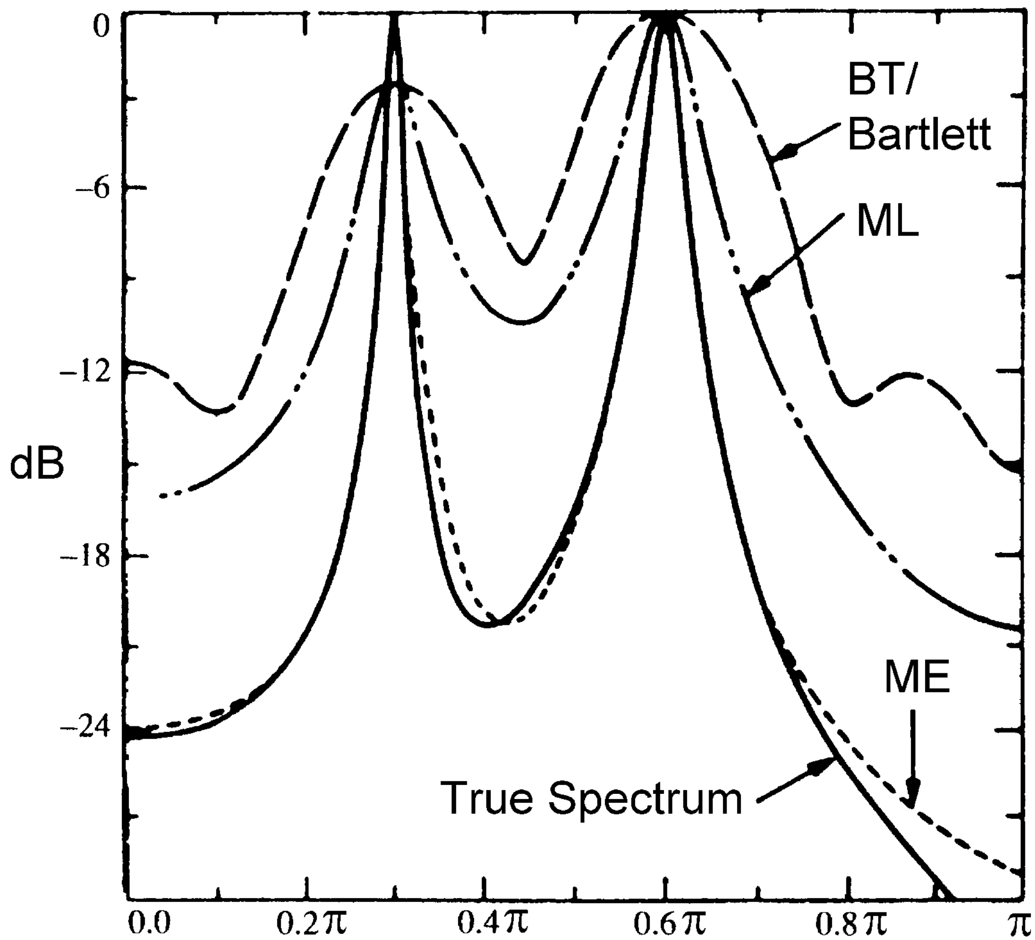
$$= \mathbf{w}^{*T} \begin{bmatrix} 1 & \cdots & 0 & 0 \\ a_1^{(N-1)} & \ddots & \vdots & \vdots \\ \vdots & \vdots & 1 & 0 \\ a_{N-1}^{(N-1)} & \cdots & a_1^{(1)} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_{N-1}^2} & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & \frac{1}{\sigma_1^2} & 0 \\ 0 & \cdots & 0 & \frac{1}{\sigma_0^2} \end{bmatrix} \begin{bmatrix} 1 & \cdots & \cdots & a_{N-1}^{(N-1)*} \\ \vdots & \ddots & \cdots & \cdots \\ 0 & \cdots & 1 & a_1^{(1)*} \\ 0 & \cdots & 0 & 1 \end{bmatrix} \mathbf{w}$$

$$= \sum_{p=0}^{N-1} \frac{\left| \sum_{k=0}^{p} a_k^{(p)} e^{-\jmath\omega k} \right|^2}{\sigma_p^2}$$

Then . . .

$$\boxed{\frac{1}{\widehat{S}_{ML}(e^{\jmath\omega})} = \sum_{p=0}^{N-1} \frac{1}{\widehat{S}_{ME}^{(p)}(e^{\jmath\omega})}}$$

SPECTRUM ANALYSIS: COMPARISON



$R_x[l] =$

$e^{-0.02l}\left(\cos 0.3\pi l + \frac{1}{15\pi}\sin 0.3\pi l\right)$

$+2e^{-0.04l}\left(\cos 0.6\pi l + \frac{1}{15\pi}\sin 0.6\pi l\right)$

(11 samples)

COMPUTATION OF THE ML ESTIMATE

FASTEST METHOD:

1. Express the denominator as

$$\mathbf{w}^{*T}\hat{\mathbf{R}}_{\boldsymbol{x}}^{-1}\mathbf{w} = \sum_{k=-N}^{N} \varrho[k] e^{-\jmath\omega k}$$

   where $\varrho[k]$ is the sum of terms on diagonals of $\hat{\mathbf{R}}_{\boldsymbol{x}}^{-1}$

2. Use the FFT to compute this term and take reciprocal

SUBSPACE METHODS

- Used for estimating *discrete* components in the spectrum of

$$x[n] = A_1 e^{\jmath\omega_1 n} + A_2 e^{\jmath\omega_2 n} + \cdots + A_M e^{\jmath\omega_M n} + \eta[n]$$

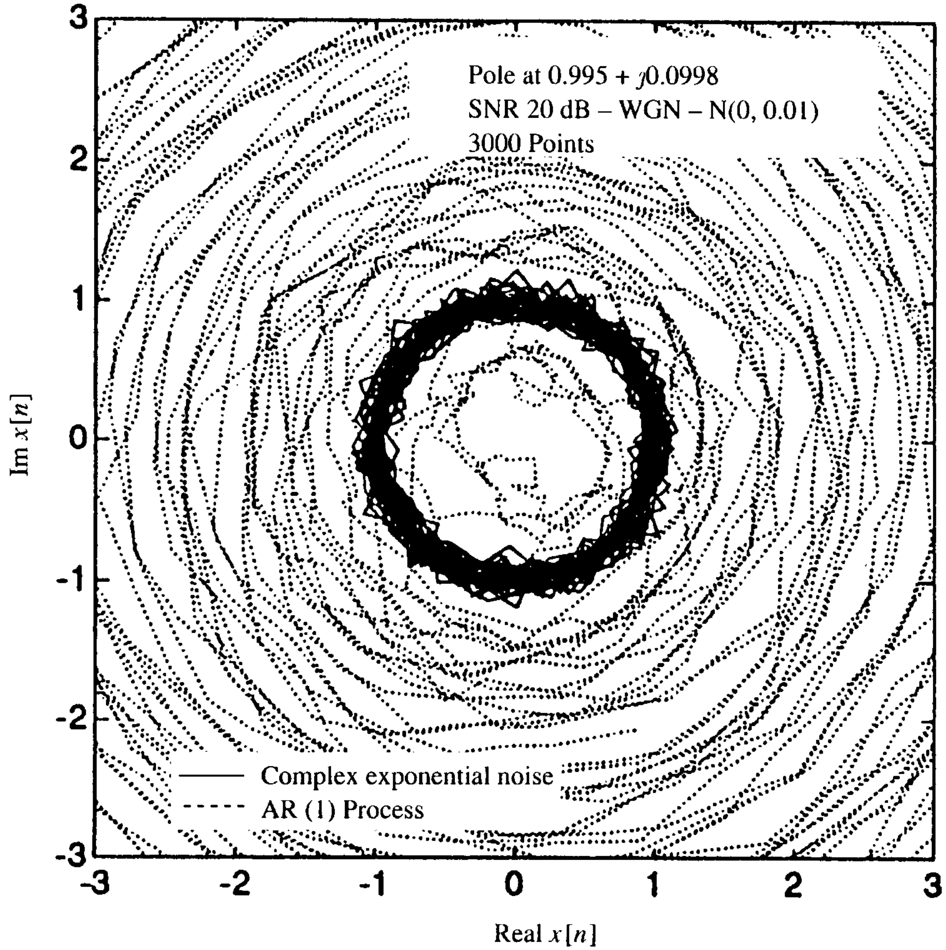
where $\eta[n]$ is white or colored "noise"

- Estimate *parameters* ω_i and $\mathsf{P}_i = \mathcal{E}\left\{|A_i|^2\right\} \quad i = 1, 2, \ldots, M$

- Based on the concept that "signals" $s_i[n] = A_i e^{\jmath\omega_i n}$ span a *subspace* of the vector space of observations

- The conditions $\mathcal{E}\left\{A_i A_k^*\right\} = \mathcal{E}\{A_i \eta^*[n]\} = 0 \quad (i \neq k)$ are assumed throughout

SUBSPACE AND AR MODELS COMPARED



SUBSPACE MODEL

$s[n] = as[n-1]$

$x[n] = s[n] + w[n]$

AR MODEL

$s[n] = as[n-1] + w[n]$

$x[n] = s[n]$

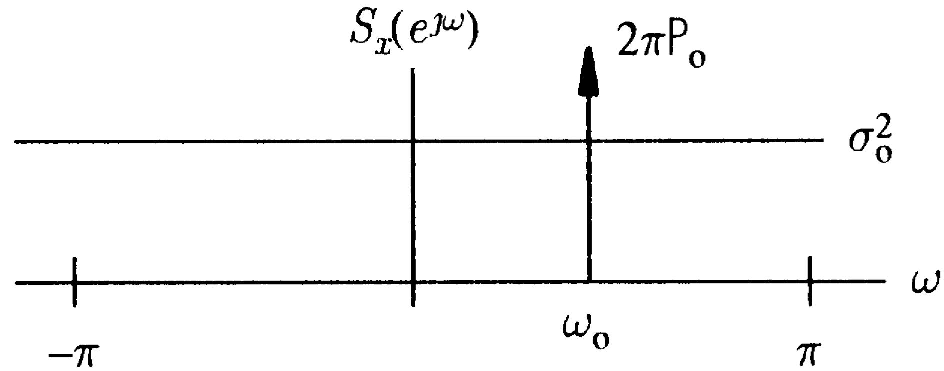
SIMPLEST PROBLEM

SIGNAL IN WHITE NOISE

$x[n] = As[n] + \eta[n]$

where $\quad s[n] = e^{\jmath\omega_o n}$

and $\quad A = |A|e^{\jmath\phi}$



VECTOR FORM AND CORRELATION MATRIX

$$\boldsymbol{x} = A\mathrm{s} + \boldsymbol{\eta} \qquad \left(\boldsymbol{x} = [x[0]\ x[1]\ \cdots x[N-1]]^T\right)$$

$$\begin{aligned} \mathbf{R}_{\boldsymbol{x}} &= \mathcal{E}\left\{A\mathrm{s}(A\mathrm{s})^{*T}\right\} + \mathcal{E}\left\{\boldsymbol{\eta}\boldsymbol{\eta}^{*T}\right\} \\ &= \mathsf{P}_\mathsf{o}\,\mathrm{s}\mathrm{s}^{*T} + \sigma_\mathsf{o}^2\,\mathbf{I} \qquad \text{where} \quad \mathsf{P}_\mathsf{o} = \mathcal{E}\left\{|A|^2\right\} \end{aligned}$$

SIMPLEST PROBLEM (cont'd.)

Observe that

- The signal vector is an eigenvector of $\mathbf{R}_{\boldsymbol{x}}$:

$$\mathbf{R}_{\boldsymbol{x}}\,\mathbf{s} = (\mathsf{P_O}\,\mathbf{s}\mathbf{s}^{*T} + \sigma_{\mathsf{O}}^{2}\,\mathbf{I})\mathbf{s} = \mathsf{P_O}\,\mathbf{s}\mathbf{s}^{*T}\mathbf{s} + \sigma_{\mathsf{O}}^{2}\,\mathbf{s} = (N\mathsf{P_O} + \sigma_{\mathsf{O}}^{2}\,)\mathbf{s}$$

- All other eigenvectors have eigenvalues equal to σ_{O}^{2}:

$$\mathbf{R}_{\boldsymbol{x}}\,\mathbf{e}_i = \mathsf{P_O}\,\mathbf{s}\,\underbrace{\mathbf{s}^{*T}\mathbf{e}_i}_{\mathbf{0}} + \sigma_{\mathsf{O}}^{2}\,\mathbf{e}_i = \sigma_{\mathsf{O}}^{2}\,\mathbf{e}_i$$

SIMPLEST PROBLEM: SOLUTION

1. Form the correlation matrix and compute its eigenvalues and eigenvectors.

2. Identify the $N-1$ smallest eigenvalues. These all have the same value, σ_{o}^2.

3. Identify the remaining (largest) eigenvalue. It is equal to $N\mathsf{P}_{\mathrm{o}} + \sigma_{\mathrm{o}}^2$. Knowledge of its value and σ_{o}^2 determines P_{o}.

4. The eigenvector corresponding to the largest eigenvalue is proportional to $\mathbf{s} = [\, 1 \;\; e^{\jmath\omega_{\mathrm{o}}} \;\; e^{\jmath 2\omega_{\mathrm{o}}} \;\; \ldots \;\; e^{\jmath(N-1)\omega_{\mathrm{o}}} \,]^T$. This in principle determines ω_{o}.

TWO SIGNALS IN WHITE NOISE

OBSERVED SEQUENCE

$$x[n] = A_1 s_1[n] + A_2 s_2[n] + \eta[n]$$

where $\quad s_i[n] = e^{\jmath \omega_i n}$ and $A_i = |A_i| e^{\jmath \phi_i}$

VECTOR FORM AND CORRELATION MATRIX

$$\boldsymbol{x} = A_1 \mathbf{s}_1 + A_2 \mathbf{s}_2 + \boldsymbol{\eta}$$

$$\mathbf{R}_{\boldsymbol{x}} = \mathsf{P}_1 \mathbf{s}_1 \mathbf{s}_1^{*T} + \mathsf{P}_2 \mathbf{s}_2 \mathbf{s}_2^{*T} + \sigma_o^2 \mathbf{I} \qquad \text{where} \quad \mathsf{P}_i = \mathcal{E}\left\{|A_i|^2\right\}$$

TWO SIGNALS IN WHITE NOISE (cont'd.)

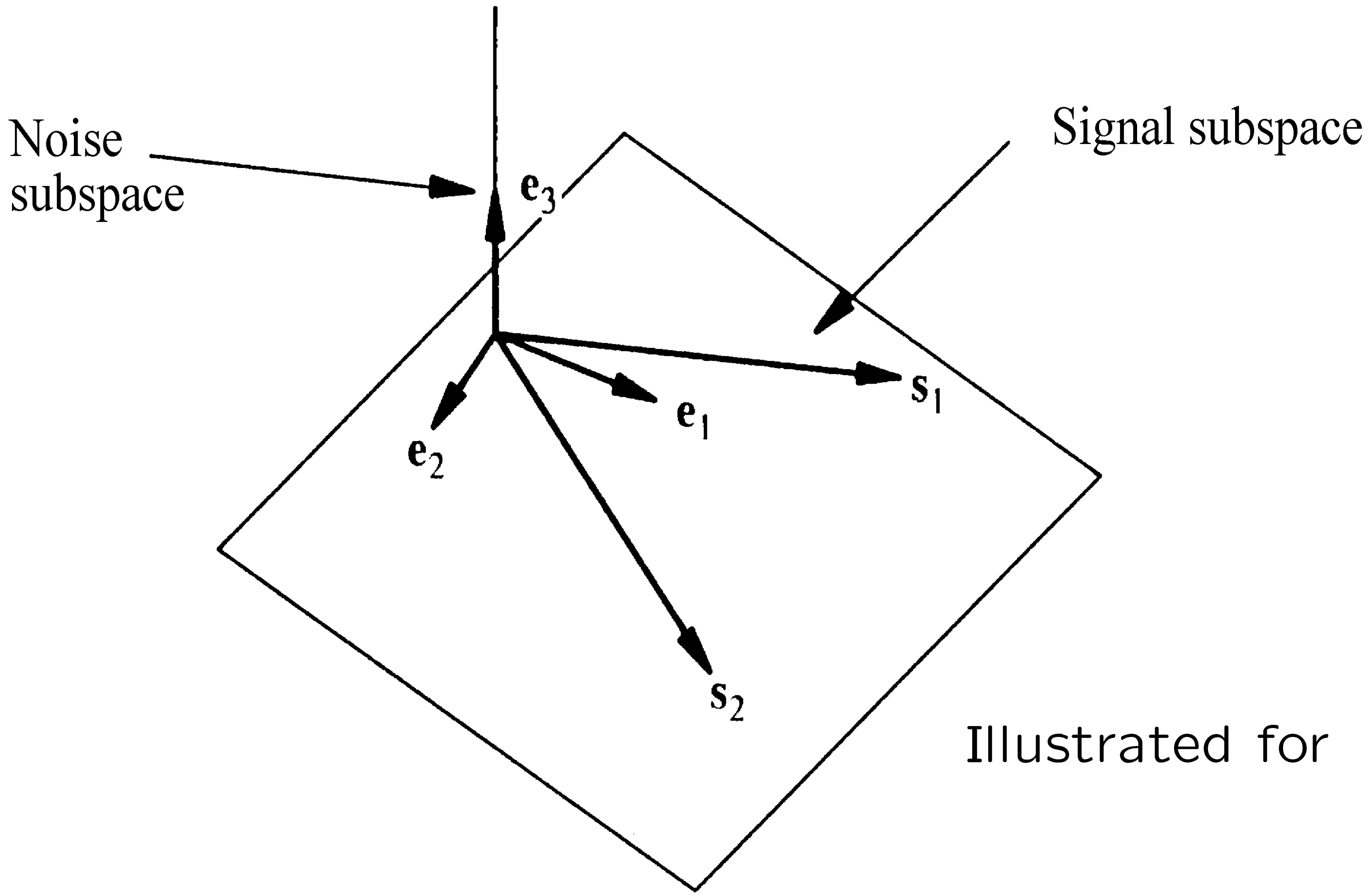
- It is possible to find $N-2$ eigenvectors orthogonal to *both* s_1 and s_2. These all have eigenvalues equal to σ_o^2:

$$\mathbf{R}_{\boldsymbol{x}}\mathbf{e}_i = \mathsf{P}_1\mathbf{s}_1 \underbrace{\mathbf{s}_1^{*T}\mathbf{e}_i}_{\mathbf{0}} + \mathsf{P}_2\mathbf{s}_2 \underbrace{\mathbf{s}_2^{*T}\mathbf{e}_i}_{\mathbf{0}} + \sigma_o^2\mathbf{I}\mathbf{e}_i = \sigma_o^2\,\mathbf{e}_i$$

- Remaining two eigenvectors lie *in* the subspace spanned by s_1 and s_2.

- Subspace spanned by s_1 and s_2 is the *signal subspace*. Complementary subspace is called the *noise subspace*.

SIGNAL AND NOISE SUBSPACES



Illustrated for $N = 3$

GENERAL PROBLEM FORMULATION: M SIGNALS IN WHITE NOISE

OBSERVED SEQUENCE AND VECTOR FORM

$$x[n] = \sum_{i=1}^{M} A_i s_i[n] + \eta[n] \qquad \Longleftrightarrow \qquad \boldsymbol{x} = \sum_{i=1}^{M} A_i \mathbf{s}_i + \boldsymbol{\eta}$$

CORRELATION MATRIX

$$\mathbf{R}_{\boldsymbol{x}} = \sum_{i=1}^{M} \mathsf{P}_i\, \mathbf{s}_i \mathbf{s}_i^{*T} + \sigma_o^2\, \mathbf{I} \qquad \text{or} \qquad \mathbf{R}_{\boldsymbol{x}} = \mathbf{S}\mathbf{P}_o\mathbf{S}^{*T} + \sigma_o^2\, \mathbf{I}$$

where

$$\mathbf{S} = \begin{bmatrix} | & | & & | \\ \mathbf{s}_1 & \mathbf{s}_2 & \cdots & \mathbf{s}_M \\ | & | & & | \end{bmatrix} \qquad \text{and} \qquad \mathbf{P}_o = \begin{bmatrix} \mathsf{P}_1 & 0 & \cdots & 0 \\ 0 & \mathsf{P}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathsf{P}_M \end{bmatrix}$$

RESULTS FOR THE GENERAL PROBLEM

- The M signal vectors $\mathbf{s}_1, \ldots, \mathbf{s}_M$ define the *signal subspace*.

- The first M eigenvectors of $\mathbf{R}_{\boldsymbol{x}}$ (corresponding to the *largest* eigenvalues) span the signal subspace. These eigenvectors have eigenvalues $> \sigma_o^2$.

- The remaining $N-M$ eigenvectors define the *noise subspace*. These all have eigenvalues equal to σ_o^2.

- The signal and noise subspaces are orthogonal and complementary.

SIGNALS IN COLORED NOISE

OBSERVATION VECTOR AND CORRELATION MATRIX

$$\boldsymbol{x} = \sum_{i=1}^{M} A_i \mathrm{s}_i + \boldsymbol{\eta} \qquad\qquad \mathbf{R}_{\boldsymbol{x}} = \mathbf{S}\mathbf{P}_{\mathsf{O}}\mathbf{S}^{*T} + \sigma_{\mathsf{O}}^2\,\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$$

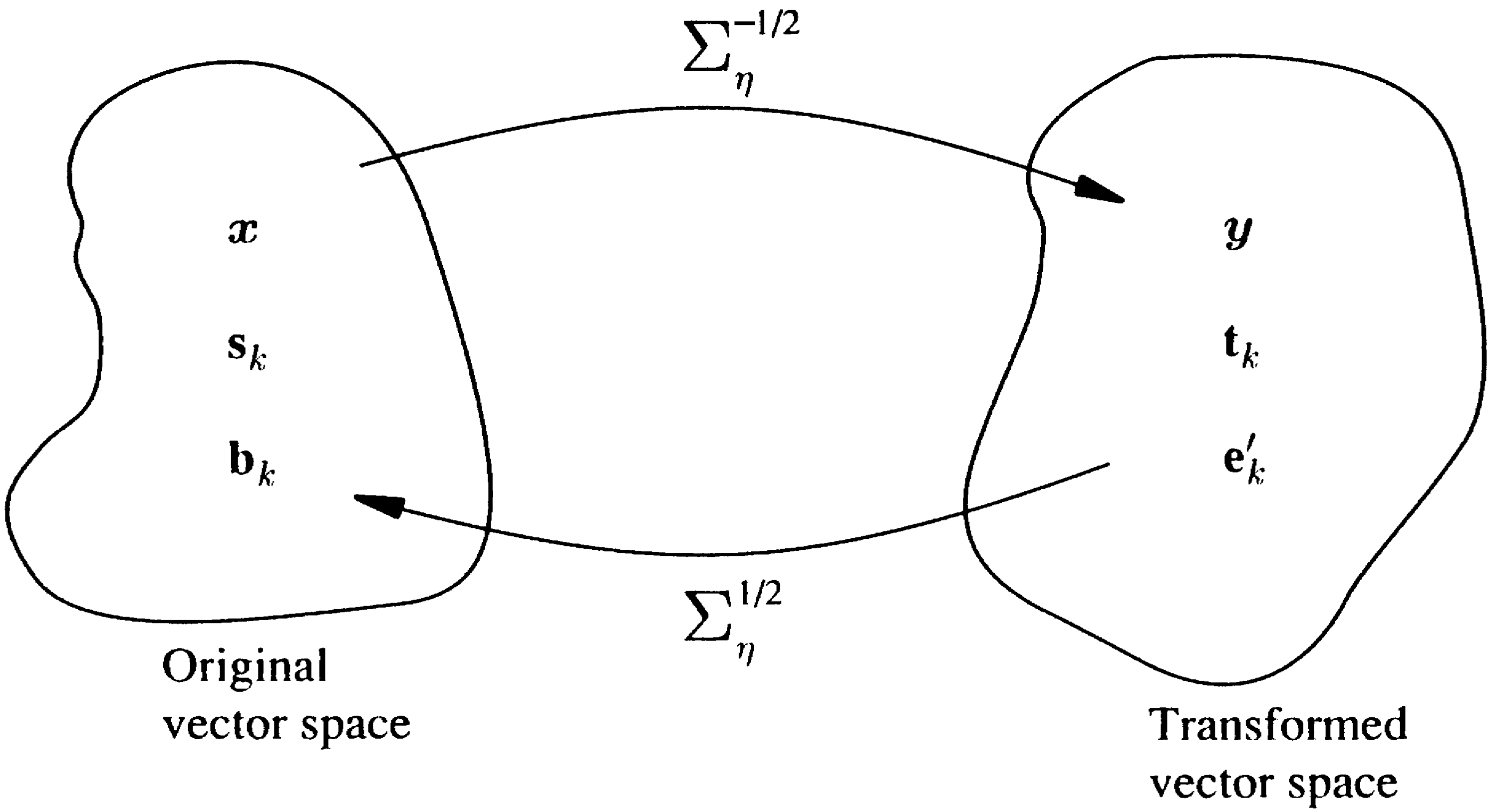
MAHALANOBIS WHITENING TRANSFORMATION

$$\boldsymbol{y} = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1/2}\boldsymbol{x} \quad \Longrightarrow \quad \mathbf{R}_{\boldsymbol{y}} = \mathbf{T}\mathbf{P}_{\mathsf{O}}\mathbf{T}^{*T} + \sigma_{\mathsf{O}}^2\,\mathbf{I}\,; \quad \mathbf{T} = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1/2}\mathbf{S}$$

Signal and noise subspace eigenvectors $\mathbf{e}'_k$ in the transformed space are represented in the original space by

$$\mathbf{b}_k = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{1/2}\mathbf{e}'_k$$

WHITENING TRANSFORMATION



EIGENVALUE PROBLEM: COLORED NOISE

- The eigenvalue problem $\mathbf{R}_{\boldsymbol{y}}\mathbf{e}'_k = \left(\boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1/2}\mathbf{R}_{\boldsymbol{x}}\boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1/2}\right)\mathbf{e}'_k = \lambda_k\mathbf{e}'_k$ in the transformed space can be replaced by a *generalized eigenvalue problem* in the original space

$$\boxed{\mathbf{R}_{\boldsymbol{x}}\mathbf{e}_k = \lambda_k\boldsymbol{\Sigma}_{\boldsymbol{\eta}}\mathbf{e}_k} \qquad \left(\text{where} \quad \mathbf{e}_k = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}^{-1/2}\mathbf{e}'_k\right)$$

- The signal and noise subspaces are then spanned by the basis vectors $\{\mathbf{b}_1, \ldots, \mathbf{b}_M\}$ and $\{\mathbf{b}_{M+1}, \ldots, \mathbf{b}_N\}$ respectively, where

$$\mathbf{b}_k = \boldsymbol{\Sigma}_{\boldsymbol{\eta}}\,\mathbf{e}_k\,, \quad k = 1, 2, \ldots, N$$

MATRICES RELATED TO SUBSPACES

EIGENVECTOR MATRICES

$$\mathbf{E}_{sig} = \begin{bmatrix} | & | & & | \\ \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_M \\ | & | & & | \end{bmatrix} \qquad \mathbf{E}_{noise} = \begin{bmatrix} | & | & & | \\ \mathbf{e}_{M+1} & \mathbf{e}_{M+2} & \cdots & \mathbf{e}_N \\ | & | & & | \end{bmatrix}$$

EIGENVALUE MATRICES

$$\mathbf{\Lambda}_{sig} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_M \end{bmatrix} \qquad \mathbf{\Lambda}_{noise} = \begin{bmatrix} \sigma_o^2 & 0 & \cdots & 0 \\ 0 & \sigma_o^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_o^2 \end{bmatrix}$$

VARIOUS IMPORTANT RELATIONS

EIGENVECTOR AND EIGENVALUE MATRICES

$$\mathbf{E} = \begin{bmatrix} \mathbf{E}_{sig} & \mathbf{E}_{noise} \end{bmatrix} \qquad \mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_{sig} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Lambda}_{noise} \end{bmatrix}$$

CORRELATION MATRIX AND INVERSE

$$\begin{aligned} \mathrm{R}_{\boldsymbol{x}} &= \mathbf{E}_{sig}\mathbf{\Lambda}_{sig}\mathbf{E}_{sig}^{*T} + \mathbf{E}_{noise}\mathbf{\Lambda}_{noise}\mathbf{E}_{noise}^{*T} \\ \mathrm{R}_{\boldsymbol{x}}^{-1} &= \mathbf{E}_{sig}\mathbf{\Lambda}_{sig}^{-1}\mathbf{E}_{sig}^{*T} + \mathbf{E}_{noise}\mathbf{\Lambda}_{noise}^{-1}\mathbf{E}_{noise}^{*T} \end{aligned}$$

PROJECTION MATRICES

$$\boldsymbol{P}_{sig} = \mathbf{E}_{sig}\mathbf{E}_{sig}^{*T} \qquad \boldsymbol{P}_{noise} = \mathbf{E}_{noise}\mathbf{E}_{noise}^{*T} = \mathbf{I} - \boldsymbol{P}_{sig}$$

PARTICULAR SUBSPACE METHODS

- Pisarenko Harmonic Decomposition
- MUSIC
- Minimum Norm
- Principal Components Linear Prediction
- *ESPRIT*

PISARENKO HARMONIC DECOMPOSITION

Assume M signals with unknown frequencies $\omega_1, \omega_2, \ldots, \omega_M$.

Take $N = M + 1$ and note that $\mathbf{e}_N$ is orthogonal to each signal vector $\mathbf{s}_i$.

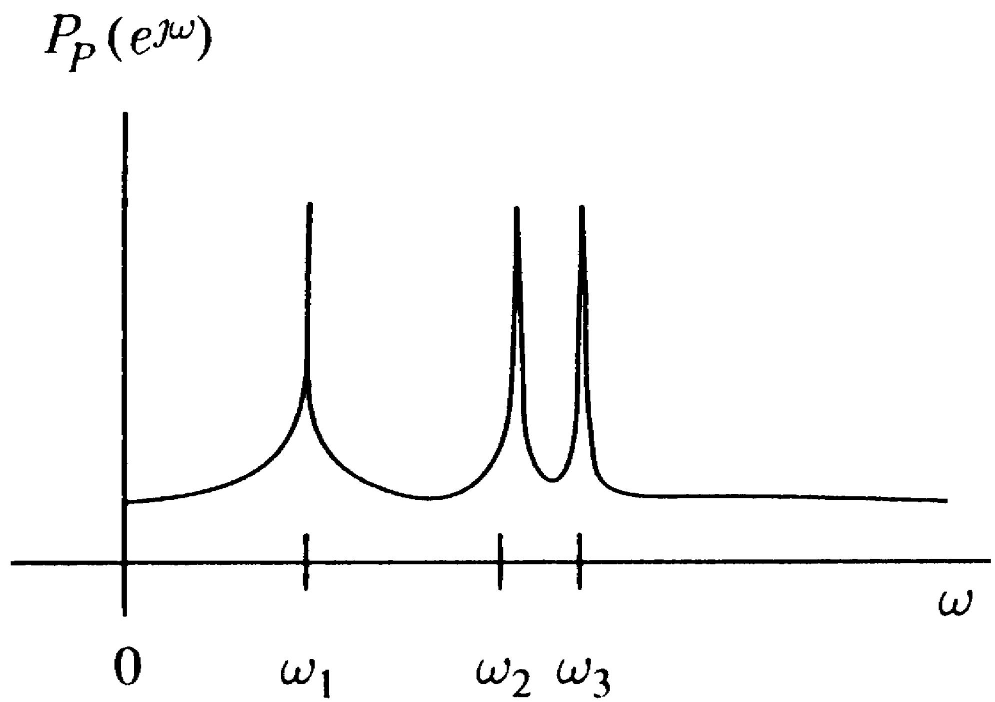
Since $\mathbf{w} = \begin{bmatrix} 1 \\ e^{\jmath\omega} \\ e^{\jmath 2\omega} \\ \vdots \\ e^{\jmath(N-1)\omega} \end{bmatrix}$ while $\mathbf{s}_i = \begin{bmatrix} 1 \\ e^{\jmath\omega_i} \\ e^{\jmath 2\omega_i} \\ \vdots \\ e^{\jmath(N-1)\omega_i} \end{bmatrix}$

it follows that

$$\mathbf{w}^{*T}\mathbf{e}_N\Big|_{\omega=\omega_i} = 0\,; \qquad i = 1, 2, \ldots, M$$

PISARENKO PSEUDOSPECTRUM

$$\widehat{P}_P(e^{\jmath\omega}) = \frac{1}{|\mathbf{w}^{*T}\mathbf{e}_N|^2} = \frac{1}{\mathbf{w}^{*T}\mathbf{e}_N\mathbf{e}_N^{*T}\mathbf{w}}$$



This function peaks at the signal frequencies.

PISARENKO - ROOT METHOD

Define the *eigenfilter*

$$E_N(z) = e_N[0] + e_N[1]z^{-1} + \cdots + e_N[N-1]z^{-(N-1)}$$

where the $e_N[n]$ are components of the eigenvector $\mathbf{e}_N$.

Then

$$E_N(e^{\jmath\omega}) = \mathbf{w}^{*T}\mathbf{e}_N$$

which goes to zero for $\omega = \omega_1, \omega_2, \ldots, \omega_M$. Therefore . . .

PISARENKO - ROOT METHOD (cont'd.)

- The M roots of $E_N(z)$ occuring on the unit circle correspond to the signal frequencies $\omega_1, \omega_2, \ldots, \omega_M$.

- The pseudospectrum can also be written in terms of $E_N(z)$ as

$$\hat{P}_P(e^{\jmath\omega}) = \frac{1}{|E_N(e^{\jmath\omega})|^2} = \frac{1}{E_N(e^{\jmath\omega})E_N^*(e^{\jmath\omega})}$$

SIGNAL POWER ESTIMATION ($M = 2$)

Write

$$\begin{aligned} \mathbf{e}_1^{*T}\mathbf{R}_{\boldsymbol{x}}\mathbf{e}_1 = \mathsf{P}_1\mathbf{e}_1^{*T}\mathbf{s}_1\mathbf{s}_1^{*T}\mathbf{e}_1 + \mathsf{P}_2\mathbf{e}_1^{*T}\mathbf{s}_2\mathbf{s}_2^{*T}\mathbf{e}_1 + \sigma_o^2 &= \lambda_1 \\ \mathbf{e}_2^{*T}\mathbf{R}_{\boldsymbol{x}}\mathbf{e}_2 = \mathsf{P}_1\mathbf{e}_2^{*T}\mathbf{s}_1\mathbf{s}_1^{*T}\mathbf{e}_2 + \mathsf{P}_2\mathbf{e}_2^{*T}\mathbf{s}_2\mathbf{s}_2^{*T}\mathbf{e}_2 + \sigma_o^2 &= \lambda_2 \end{aligned}$$

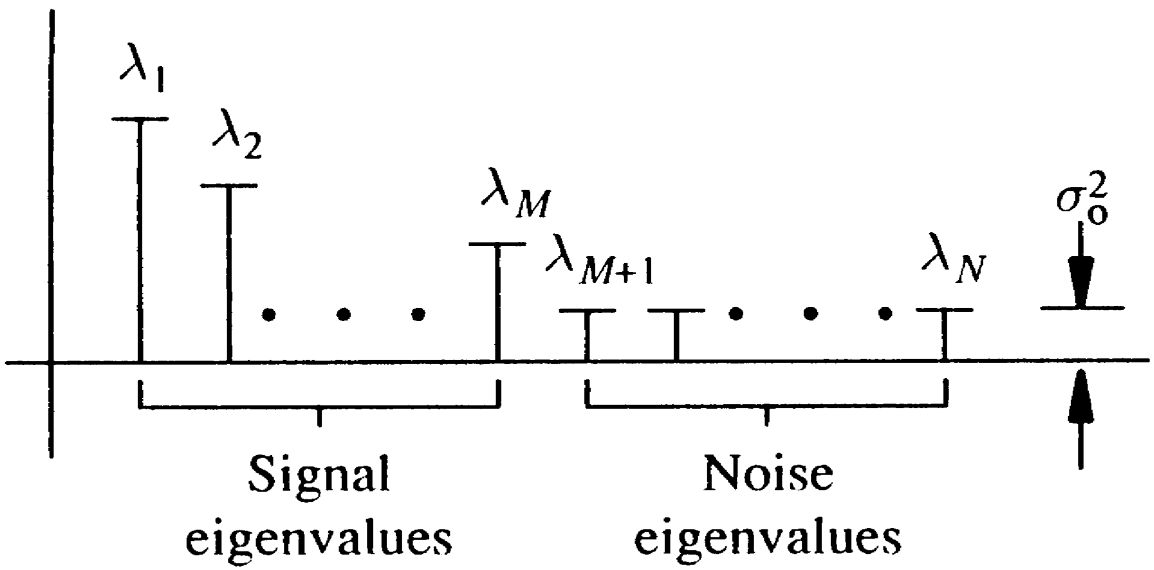
These are linear equations of the form

$$\begin{bmatrix} |\beta_{11}|^2 & |\beta_{12}|^2 \\ |\beta_{21}|^2 & |\beta_{22}|^2 \end{bmatrix} \begin{bmatrix} \mathsf{P}_1 \\ \mathsf{P}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 - \sigma_o^2 \\ \lambda_2 - \sigma_o^2 \end{bmatrix} \qquad \text{where} \qquad \beta_{ik} = \mathbf{e}_i^{*T}\mathbf{s}_k$$

These can be solved for P_1 and P_2.

MUSIC (MULTIPLE SIGNAL CLASSIFICATION)

- Uses correlation matrix of any size $N > M + 1$
- Can be used to estimate the number of signals M



MUSIC: FREQUENCY ESTIMATION

PSEUDOSPECTRUM

$$\boxed{\hat{P}_{MU}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}\boldsymbol{P}_{noise}\mathbf{w}} = \frac{1}{\mathbf{w}^{*T}\mathbf{E}_{noise}\mathbf{E}_{noise}^{*T}\mathbf{w}} = \frac{1}{\sum_{i=M+1}^{N}|E_i(e^{\jmath\omega})|^2}}$$

ROOT METHOD (ROOT MUSIC)

Find roots of polynomial

$$\hat{P}_{MU}^{-1}(z) = \sum_{i=M+1}^{N} E_i(z)E_i^*(1/z^*)$$

lying *on* the unit circle, where $E_i(z)$ is an eigenfilter.

Remaining roots are called "spurious."

MUSIC VARIATION

An alternative pseudospectrum can be defined as

$$\widehat{P}'_{MU}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}(\sum_{i=M+1}^{N} \frac{1}{\lambda_i}\mathbf{e}_i\mathbf{e}_i^{*T})\mathbf{w}}$$

In theory, this is equivalent to $\frac{\sigma_o^2}{\mathbf{w}^{*T}(\sum_{i=M+1}^{N} \mathbf{e}_i\mathbf{e}_i^{*T})\mathbf{w}}$ which differs from the regular MUSIC pseudospectrum by only a constant.

In practice, however, by using the estimated eigenvalues $\lambda_M, \ldots \lambda_N$ the performance is sometimes inproved.

COMPARISON OF METHODS

Maximum Likelihood	$\widehat{S}_{ML}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}\mathbf{R}_{\boldsymbol{x}}^{-1}\mathbf{w}}$
Maximum Entropy	$\widehat{S}_{ME}(e^{\jmath\omega}) = \frac{\sigma_{N-1}^2}{\mathbf{w}^{*T}\mathbf{a}_{N-1}\mathbf{a}_{N-1}^{*T}\mathbf{w}}$
MUSIC	$\widehat{P}'_{MU}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}\sum_{i=M+1}^{N}\frac{1}{\lambda_i}\mathbf{e}_i\mathbf{e}_i^{*T}\mathbf{w}}$

- All represent different decompositions of the inverse correlation matrix!

MINIMUM-NORM PROCEDURE

- The frequency vector $\mathbf{w}$ is projected onto a *single vector* $\mathbf{d}$ lying in the *noise subspace*. The vector $\mathbf{d}$ is chosen to have *minimum norm* $\|\mathbf{d}\|$ subject to the constraint $d[0] = 1$.

- If the noise subspace eigenvector matrix is partitioned as

$$\mathbf{E}_{noise} = \begin{bmatrix} \mathbf{c}^{*T} \\ \\ \mathbf{E}'_{noise} \end{bmatrix} \quad \text{then} \quad \mathbf{d} = \begin{bmatrix} 1 \\ \\ \mathbf{E}'_{noise}\mathbf{c}/(\mathbf{c}^{*T}\mathbf{c}) \end{bmatrix}$$

- $\mathbf{d}$ can also be interpreted as the *total least squares* solution to the *linear prediction* problem for the data [Dowling and DeGroat, 1991].

MINIMUM-NORM: FREQUENCY ESTIMATION

PSEUDOSPECTRUM

$$\widehat{P}_{MN}(e^{\jmath\omega}) \stackrel{\text{def}}{=} \frac{1}{|\mathbf{w}^{*T}\mathbf{d}|^2} = \frac{1}{\mathbf{w}^{*T}\mathbf{d}\mathbf{d}^{*T}\mathbf{w}} = \frac{1}{|D(e^{\jmath\omega})|^2}$$

ROOT METHOD

Find roots of the polynomial

$$D(z) = \sum_{k=0}^{N-1} d[k] z^{-k}$$

lying *on* the unit circle $(d[k]$ are components of $\mathbf{d})$.

WHY MINIMUM NORM?

The polynomial $D(z)$ can be factored as

$$D(z) = D_1(z) \cdot D_2(z)$$

where

- $D_1(z)$ has roots only *on* the unit circle (due to signals)

- $D_2(z)$ has roots only *within* the unit circle (spurious roots)

In other words, $D_2(z)$ is a minimum-phase polynomial.

The roots of $D_2(z)$ are approximately uniformly distributed around the inside of the unit circle, away from the roots of $D_1(z)$.

MINIMUM-NORM SOLUTION
(GENERAL PROCEDURE)

- It is desired to minimize $\|\mathbf{d}\|^2 = \mathbf{d}^{*T}\mathbf{d}$

- Constraints are

$$\mathbf{d} = \boldsymbol{P}_{noise}\mathbf{d} = \mathbf{E}_{noise}\mathbf{E}_{noise}^{*T}\mathbf{d} \qquad \text{and} \qquad \mathbf{d}^{*T}\boldsymbol{\iota} = 1$$

- Form the Lagrangian

$$\mathcal{L} = \mathbf{d}^{*T}\mathbf{d} + \mu(1 - \mathbf{d}^{*T}\mathbf{E}_{noise}\mathbf{E}_{noise}^{*T}\boldsymbol{\iota}) + \mu^*(1 - \boldsymbol{\iota}^T\mathbf{E}_{noise}\mathbf{E}_{noise}^{*T}\mathbf{d})$$

- Set $\nabla_{\mathbf{d}^*}\mathcal{L} = \mathbf{0}$ and solve for $\mathbf{d}$ (see text for details).

PRINCIPAL COMPONENTS LINEAR PREDICTION

- Principal components methods use the *principal components* approximation to the correlation matrix, or its principal components inverse

$$\mathbf{R}_{\boldsymbol{x}}^{(M)} \stackrel{\text{def}}{=} \sum_{i=1}^{M} \lambda_i \mathbf{e}_i \mathbf{e}_i^{*T} \qquad \mathbf{R}_{\boldsymbol{x}}^{+(M)} = \sum_{i=1}^{M} \frac{1}{\lambda_i} \mathbf{e}_i \mathbf{e}_i^{*T}$$

- The Principal Components Linear Prediction method [Tufts and Kumaresan, 1982] exploits this technique to produce a highly effective procedure for the estimation of complex exponentials (or sinusoids) in noise.

LINEAR PREDICTION FOR COMPLEX EXPONENTIAL SIGNALS AND NO NOISE

CORRELATION MATRIX

$$\mathbf{R}_{\boldsymbol{x}} = \mathbf{R}_{\boldsymbol{s}} = \sum_{i=1}^{M} \mathsf{P}_i \mathbf{s}_i \mathbf{s}_i^{*T} = \mathbf{S}^{*T} \mathbf{P}_{\mathsf{O}} \mathbf{S}$$

NORMAL EQUATIONS

$$\mathbf{R}_{\boldsymbol{x}} \mathbf{a} = \mathbf{0}$$

- $\mathbf{a}$ is an eigenvector corresponding to eigenvalue $\lambda = 0$.
- $\mathbf{a}$ lies in the noise subspace ($\mathbf{s}_i^{*T} \mathbf{a} = 0$).
- The "noise subspace" is the null space of $\mathbf{R}_{\boldsymbol{x}}$.

LINEAR PREDICTION: NO NOISE (cont'd.)

- The prediction error filter suggests a pseudospectrum

$$\widehat{P}_x(e^{\jmath\omega}) = \frac{1}{|\mathbf{w}^{*T}\mathbf{a}|^2} = \frac{1}{|A(e^{\jmath\omega})|^2}$$

  which peaks at the desired frequencies.

- For $N > M + 1$ the Normal equations $\mathbf{R}_{\boldsymbol{x}}\mathbf{a} = \mathbf{0}$ have *multiple* solutions. Choose the *minimum-norm* solution.

LINEAR PREDICTION: NO NOISE (cont'd.)

MINIMUM-NORM SOLUTION

By dropping the top row of the matrix $\mathbf{R}_{\boldsymbol{x}}$ and defining

$\mathbf{a} = \begin{bmatrix} 1 \\ \\ \mathbf{a}' \end{bmatrix}$ the Normal equations $\mathbf{R}_{\boldsymbol{x}}\mathbf{a} = \mathbf{0}$ can be reduced to

$$\mathbf{R}'_{\boldsymbol{x}}\mathbf{a}' = -\mathbf{r}$$

The minimum-norm solution can then be expressed as

$$\mathbf{a}' = -\mathbf{R}'^{+}_{\boldsymbol{x}}\mathbf{r} = -\sum_{i=1}^{M} \left(\frac{\mathbf{e}'^{*T}_i \mathbf{r}}{\lambda'_i} \right) \mathbf{e}'_i$$

LINEAR PREDICTION WITH NOISE

The Normal equations $\mathbf{R}_{\boldsymbol{x}}\mathbf{a} = \sigma^2 \boldsymbol{\iota}$ have the solution

$$\mathbf{a}' = -\sum_{i=1}^{P} \left(\frac{\mathbf{e}_i'^{*T}\mathbf{r}}{\lambda_i'} \right) \mathbf{e}_i' \; ; \qquad P = N - 1$$

The PCLP method instead uses

$$\mathbf{a}' = -\sum_{i=1}^{M} \left(\frac{\mathbf{e}_i'^{*T}\mathbf{r}}{\lambda_i'} \right) \mathbf{e}_i' = -\mathbf{R}_{\boldsymbol{x}}'^{+(M)}\mathbf{r}$$

If noise power is not too large, eigenvectors $\mathbf{e}_i'$ for $i = 1, 2, \ldots, M$ are approximately the same as without noise.

PCLP: FREQUENCY ESTIMATION

PSEUDOSPECTRUM

$$\widehat{P}_{PCLP}(e^{\jmath\omega}) = \frac{1}{|\mathbf{w}^{*T}\mathbf{a}|^2} = \frac{1}{|A(e^{\jmath\omega})|^2}$$

with $\mathbf{a} = [1\ \mathbf{a}'^T]^T$ and $\mathbf{a}' = -\mathbf{R}'^{+(M)}_{\boldsymbol{x}}\mathbf{r}$.

ROOT METHOD

Find roots lying *on* the unit circle of

$$A(z) = \sum_{k=0}^{P} a_k z^{-k}$$

TEST CASE: LP VS PCLP

- Data: $\mathrm{x}[n] = e^{\jmath 1.00\pi n + \pi/4} + e^{\jmath 1.04\pi n} + \eta[n]\,;\quad N_s = 25$ samples

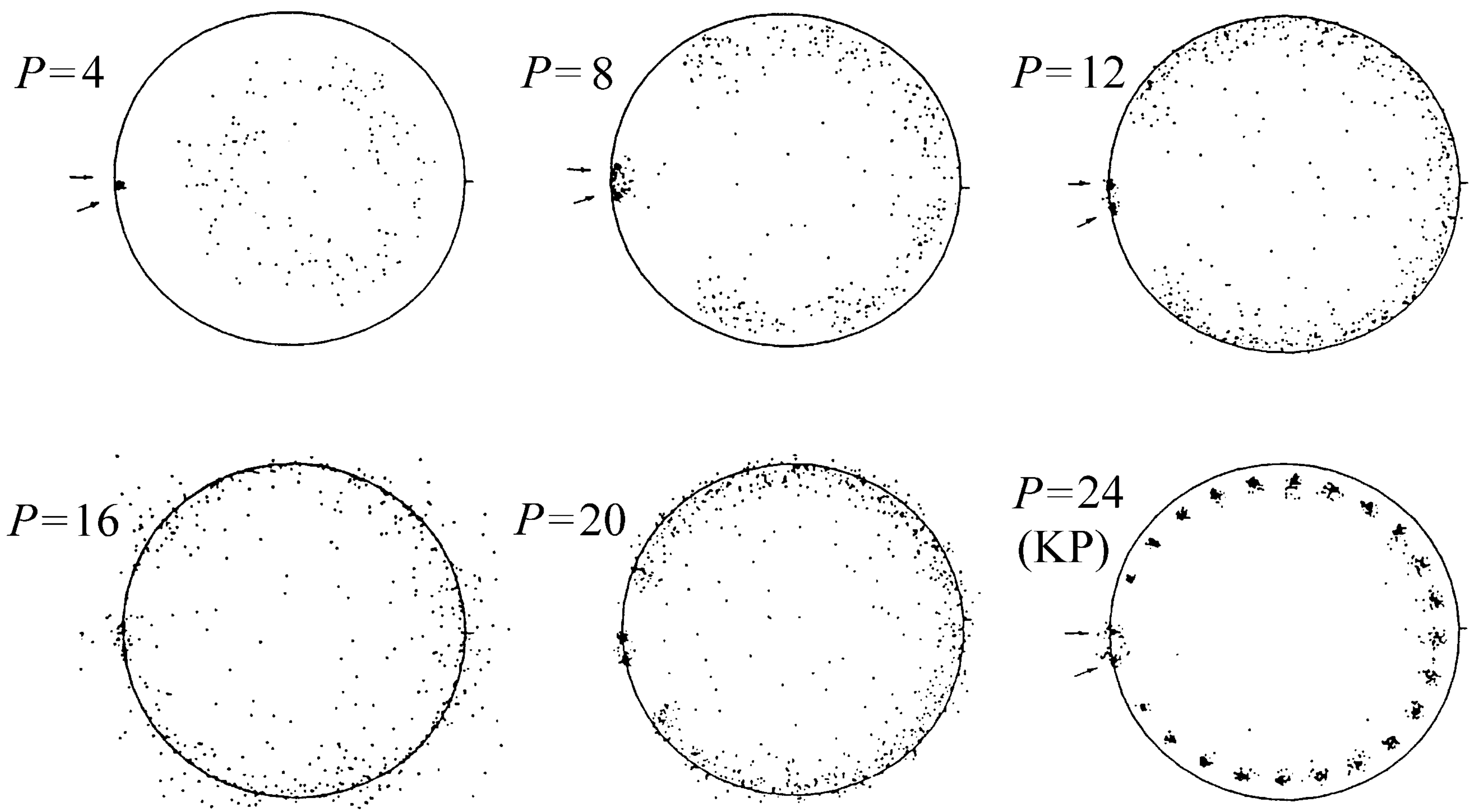
  $\eta[n]$ is white noise with SNR $= -10\log_{10}\sigma_0^2 = 10$dB.

- Roots of $A(z)$ are plotted for various prediction orders P using the *modified covariance method* (50 trials per plot).

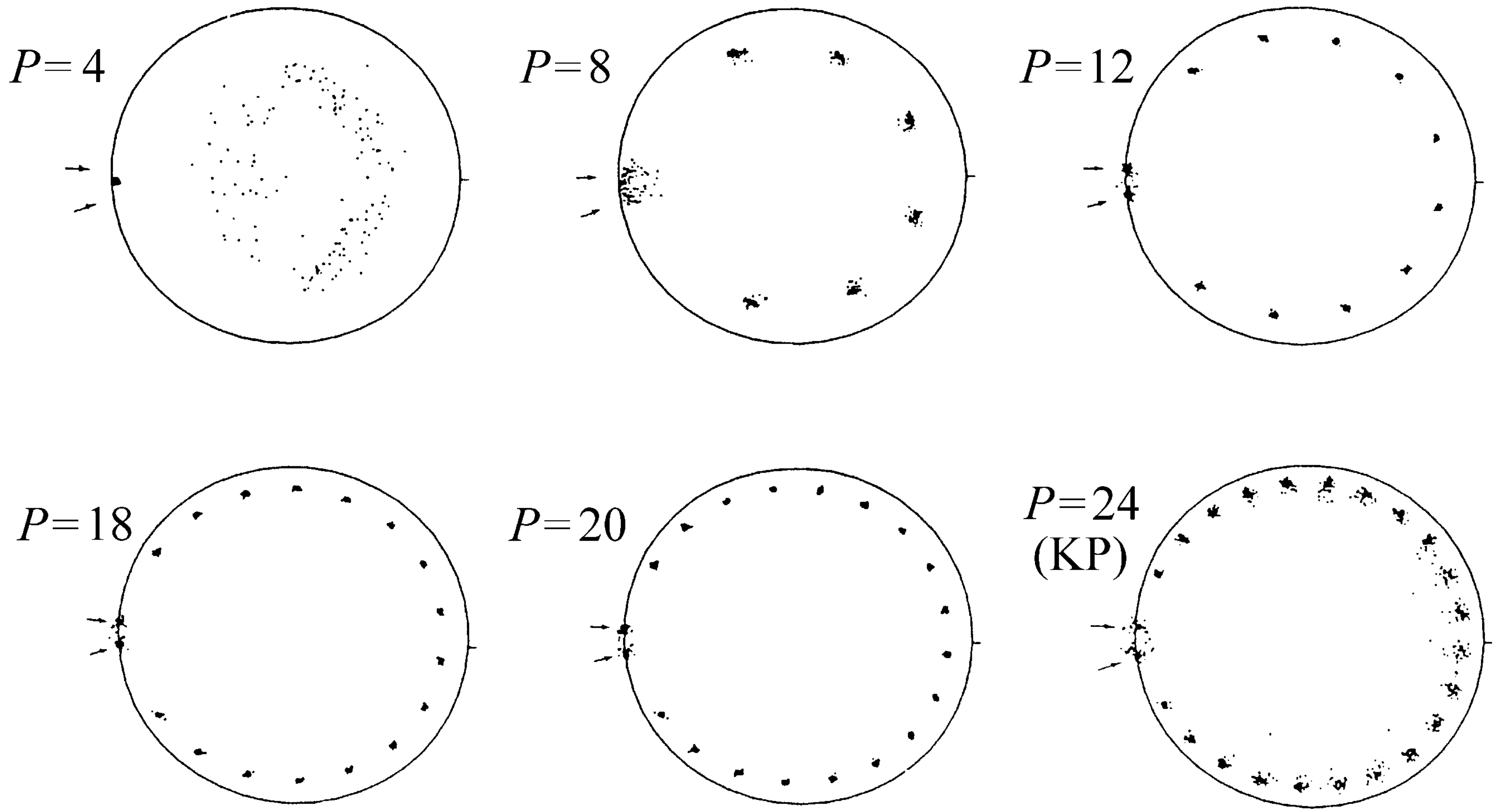
- For $P = N_s - M/2$ the rank of $\mathbf{R}'_{\boldsymbol{x}}$ is reduced to M $(= 2)$. This is called the "Kumaresan-Prony" case.

- Recommended prediction order for PCLP is $P \approx \frac{3}{4}N_s$.

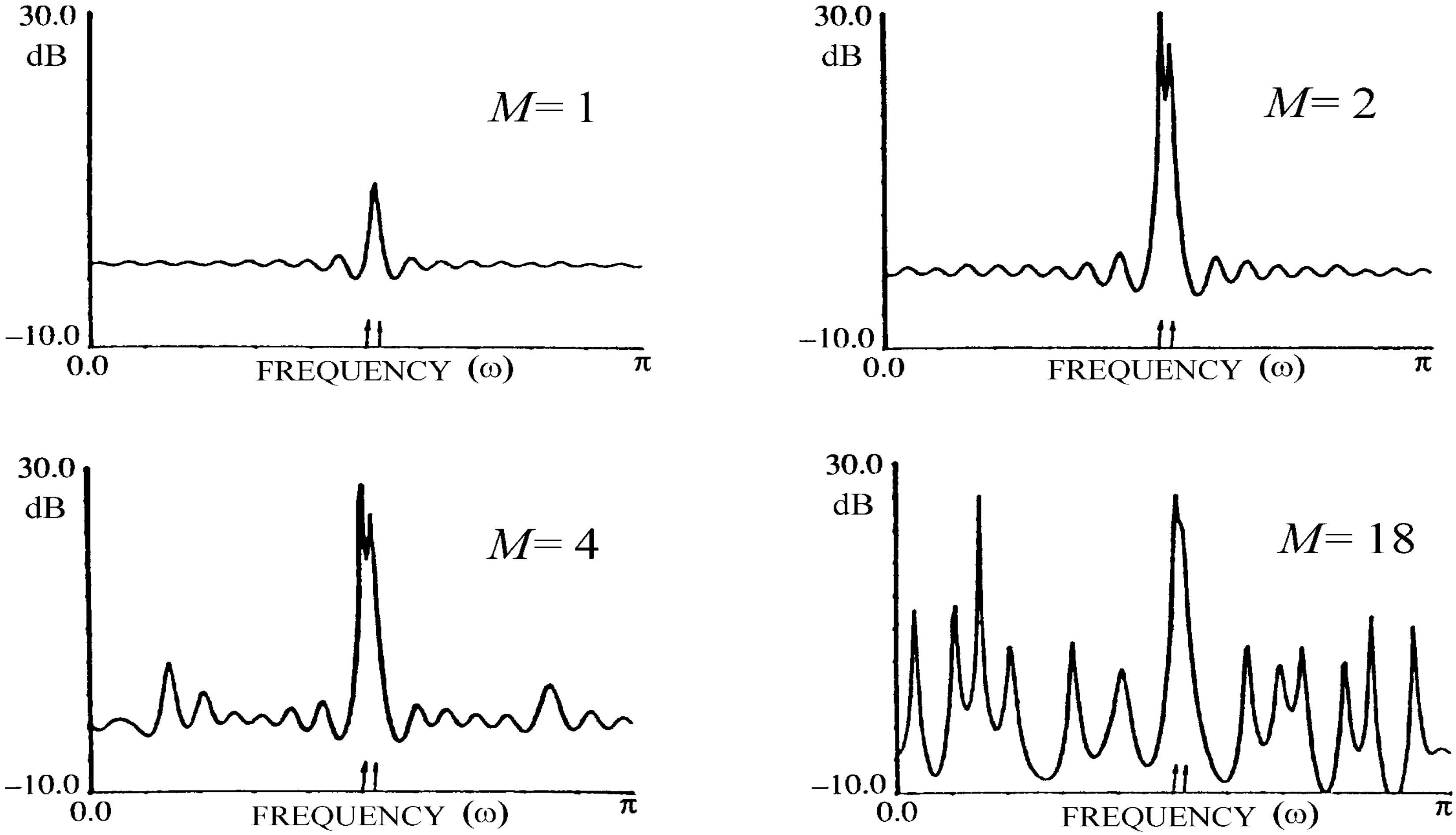
LINEAR PREDICTION RESULTS



PCLP RESULTS FOR $M = 2$



PCLP PSEUDOSPECTRA FOR $P = 18$



ESPRIT

(ESTIMATION OF SIGNAL PARAMETERS VIA ROTATIONAL INVARIANCE TECHNIQUES)

- Exploits an invariance principle that naturally exists for discrete time signals

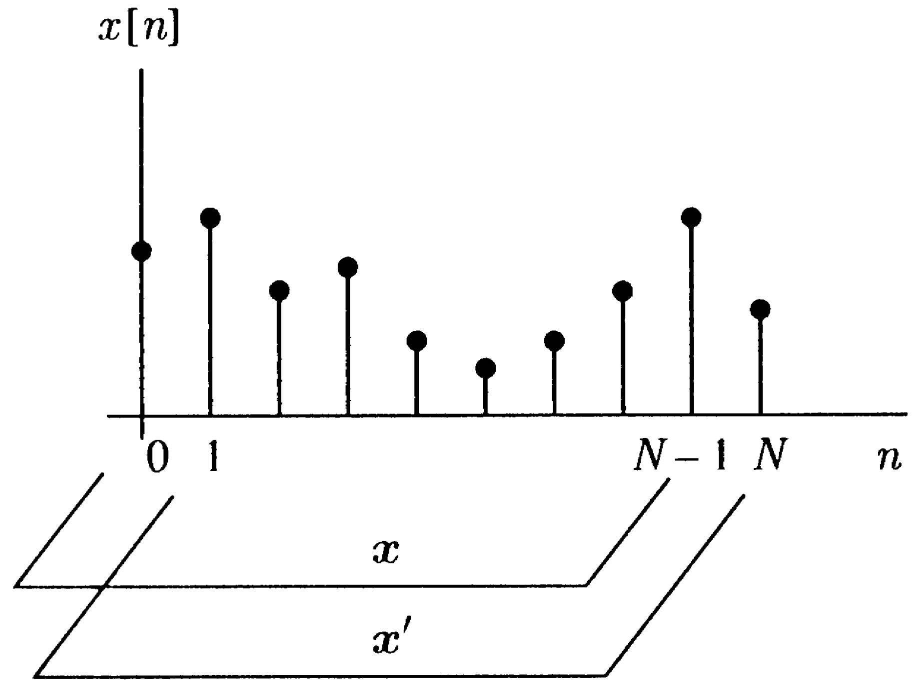
- Original technique described first to motivate the method

- Current TLS version then described

ESPRIT SIGNAL MODEL

$$\boldsymbol{x} = \sum_{i=1}^{M} A_i \mathrm{s}_i + \boldsymbol{\eta}$$

$$\boldsymbol{x}' = \sum_{i=1}^{M} A_i \mathrm{s}'_i + \boldsymbol{\eta}'$$



Note that $\mathrm{s}'_i = \left[\begin{array}{cccc} e^{\jmath\omega_i} & e^{\jmath 2\omega_i} & \cdots & e^{\jmath N\omega_i} \end{array} \right]^T = e^{\jmath\omega_i}\,\mathrm{s}_i$

ESPRIT FORMULATION

$$\mathbf{R}_{\boldsymbol{x}} = \sum_{i=1}^{M} \mathsf{P}_i \mathbf{s}_i \mathbf{s}_i^{*T} + \sigma_\mathsf{o}^2 \mathbf{I} = \mathbf{S}\mathbf{P}_\mathsf{o}\mathbf{S}^{*T} + \sigma_\mathsf{o}^2 \mathbf{I}$$

$$\mathbf{R}_{\boldsymbol{x}\boldsymbol{x}'} = \sum_{i=1}^{M} \mathsf{P}_i e^{-\jmath\omega_i} \mathbf{s}_i \mathbf{s}_i^{*T} + \sigma_\mathsf{o}^2 \mathbf{D}_{-1} = \mathbf{S}\mathbf{P}_\mathsf{o}\mathbf{\Phi}^*\mathbf{S}^{*T} + \sigma_\mathsf{o}^2 \mathbf{D}_{-1}$$

where

$$\mathbf{\Phi} = \begin{bmatrix} e^{\jmath\omega_1} & 0 & \cdots & 0 \\ 0 & e^{\jmath\omega_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\jmath\omega_M} \end{bmatrix} \qquad \mathbf{D}_{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}$$

IDENTIFYING THE FREQUENCIES

Form $\mathbf{R}_{\boldsymbol{s}} \stackrel{\text{def}}{=} \mathbf{R}_{\boldsymbol{x}} - \sigma_{\mathsf{o}}^2 \mathbf{I} = \mathbf{S}\mathbf{P}_{\mathsf{o}}\mathbf{S}^{*T}$

and $\mathbf{R}_{\boldsymbol{s}\boldsymbol{s}'} \stackrel{\text{def}}{=} \mathbf{R}_{\boldsymbol{x}\boldsymbol{x}'} - \sigma_{\mathsf{o}}^2 \mathbf{D}_{-1} = \mathbf{S}\mathbf{P}_{\mathsf{o}}\boldsymbol{\Phi}^*\mathbf{S}^{*T}$

Then consider

$$\begin{aligned}
\mathbf{R}_{\boldsymbol{s}}\check{\mathbf{e}} - \check{\lambda}\mathbf{R}_{\boldsymbol{s}\boldsymbol{s}'}\check{\mathbf{e}} &= \mathbf{S}\mathbf{P}_{\mathsf{o}}(\mathbf{I} - \check{\lambda}\boldsymbol{\Phi}^*)\mathbf{S}^{*T}\check{\mathbf{e}} \\
&= \mathbf{S}\mathbf{P}_{\mathsf{o}} \begin{bmatrix} 1 - \check{\lambda}e^{-\jmath\omega_1} & & \cdots & 0 \\ 0 & & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & & 1 - \check{\lambda}e^{-\jmath\omega_M} \end{bmatrix} \mathbf{S}^{*T}\check{\mathbf{e}} = \mathbf{0}
\end{aligned}$$

Since $\check{\lambda}_k = e^{\jmath\omega_k}$ reduces the rank of this matrix, λ_k is a *generalized eigenvalue* of

$$\mathbf{R}_{\boldsymbol{s}}\,\check{\mathbf{e}} = \check{\lambda}\mathbf{R}_{\boldsymbol{s}\boldsymbol{s}'}\,\check{\mathbf{e}}$$

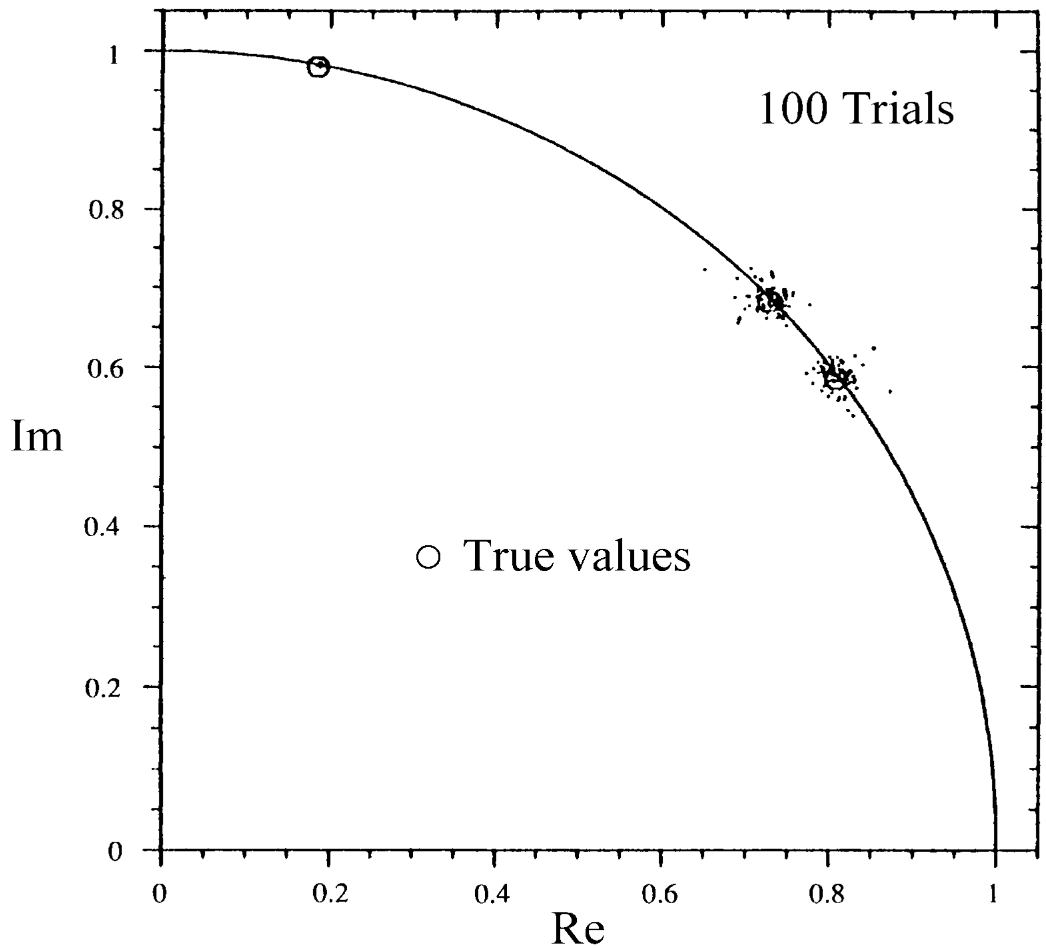
ESPRIT FREQUENCY ESTIMATE RESULTS

$N = 7$

SNR = 20 dB

Frequency values:

$$\omega_1 = 0.10\pi$$
$$\omega_2 = 0.12\pi$$
$$\omega_3 = 0.22\pi$$



PROBLEM WITH ORIGINAL FORMULATION

- The matrices $\mathbf{R}_{\boldsymbol{s}}$ and $\mathbf{R}_{\boldsymbol{ss}'}$ are not of full rank and have identical null spaces. Therefore the generalized eigenvalue problem

$$\mathbf{R}_{\boldsymbol{s}}\,\check{\mathbf{e}} = \check{\lambda}\mathbf{R}_{\boldsymbol{ss}'}\,\check{\mathbf{e}}$$

  is ill-posed.

- Since the matrices are not of full rank, the eigenvalues $\check{\lambda}$ corresponding to "eigenvectors" $\check{\mathbf{e}}$ in the null-space of these matrices can assume *any* value – and so are *not defined!*

- In practice, $\mathbf{R}_{\boldsymbol{s}}$ and $\mathbf{R}_{\boldsymbol{ss}'}$ may not have zero rank, but will be at least poorly conditioned.

TLS *ESPRIT*

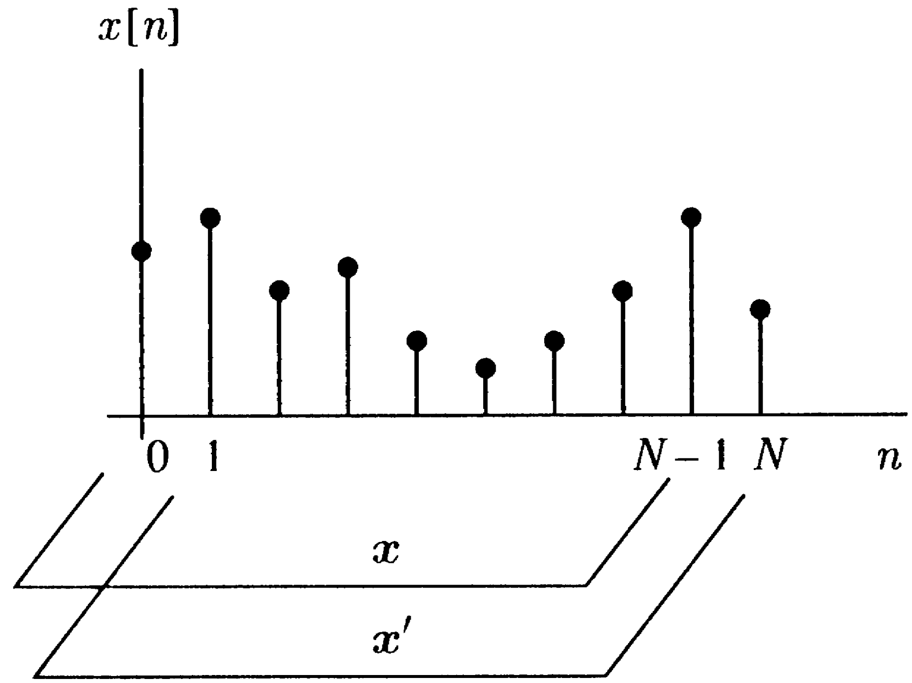
- Exploits an invariance property of the signal subspaces similar to the invariance property of $\mathbf{R}_{\boldsymbol{s}}$ and $\mathbf{R}_{\boldsymbol{ss}'}$.

- Relates to the theory of a matrix "pencil" and rank-reducing numbers.

- A least squares version of *ESPRIT* is also possible but the total least squares version is preferable.

DEFINITION OF EXPANDED VECTORS

$$\bar{\boldsymbol{x}} \stackrel{\text{def}}{=} \begin{bmatrix} \boldsymbol{x} \\ \\ x[N] \end{bmatrix} = \begin{bmatrix} x[0] \\ \\ \boldsymbol{x}' \end{bmatrix}$$



- The overbar will be used to refer to quantities relating to the expanded vectors.

APPLICATION OF INVARIANCE

The signal subspace is spanned by columns of the matrix

$$\bar{\mathbf{S}} \stackrel{\text{def}}{=} \begin{bmatrix} | & | & & | \\ \bar{\mathbf{s}}_1 & \bar{\mathbf{s}}_2 & \cdots & \bar{\mathbf{s}}_M \\ | & | & & | \end{bmatrix} = \begin{bmatrix} & \mathbf{S} & \\ & & \\ \times & \cdots & \times \end{bmatrix} = \begin{bmatrix} \times & \cdots & \times \\ & & \\ & \mathbf{S\Phi} & \end{bmatrix}$$

Any other set of basis vectors for the signal subspace

$$\bar{\mathbf{B}} = \begin{bmatrix} | & | & & | \\ \bar{\mathbf{b}}_1 & \bar{\mathbf{b}}_2 & \cdots & \bar{\mathbf{b}}_M \\ | & | & & | \end{bmatrix}$$

can be related to $\bar{\mathbf{S}}$ by a nonsingular transformation

$$\bar{\mathbf{B}}\boldsymbol{\Upsilon} = \bar{\mathbf{S}}$$

APPLICATION OF INVARIANCE (cont'd.)

The relation $\bar{\mathbf{B}}\boldsymbol{\Upsilon} = \bar{\mathbf{S}}$ can be written in two forms:

$$\bar{\mathbf{B}}\boldsymbol{\Upsilon} = \begin{bmatrix} & \mathbf{B} & \\ & & \\ \times & \cdots & \times \end{bmatrix} \boldsymbol{\Upsilon} = \begin{bmatrix} & \mathbf{S} & \\ & & \\ \times & \cdots & \times \end{bmatrix}$$

and

$$\bar{\mathbf{B}}\boldsymbol{\Upsilon} = \begin{bmatrix} \times & \cdots & \times \\ & & \\ & \mathbf{B}' & \end{bmatrix} \boldsymbol{\Upsilon} = \begin{bmatrix} \times & \cdots & \times \\ & & \\ & \mathbf{S}\boldsymbol{\Phi} & \end{bmatrix}$$

Thus

$$\mathbf{B}'\boldsymbol{\Upsilon} = \mathbf{S}\boldsymbol{\Phi} = (\mathbf{B}\boldsymbol{\Upsilon})\,\boldsymbol{\Phi} \quad \Longrightarrow \quad \mathbf{B}'\boldsymbol{\Upsilon} = \mathbf{B}\boldsymbol{\Upsilon}\boldsymbol{\Phi}$$

The last equation can be rewritten as . . .

ESPRIT FUNDAMENTAL RELATIONS

INVARIANCE OF SUBSPACES

$$\boxed{\mathrm{B}\mathbf{\Psi} = \mathrm{B}'}$$

(solve for $\mathbf{\Psi}$)

EIGEN-DECOMPOSITION OF TRANSFORMATION

$$\boxed{\mathbf{\Psi} = \mathbf{\Upsilon}\mathbf{\Phi}\mathbf{\Upsilon}^{-1}}$$

- Eigenvalues of $\mathbf{\Psi}$ have the form $e^{\jmath\omega_k}$

IMPLEMENTATION OF *ESPRIT*

- Basis vectors $\bar{\mathbf{B}}$ can be found as eigenvectors of the correlation matrix, *a la* MUSIC.

- In theory the invariance relation

$$\mathbf{B}\Psi = \mathbf{B}'$$

  is satisfied *exactly.* In practice, this is an overdetermined set of linear equations.

- Least squares solution produces LS version; total least squares solution leads to TLS version of *ESPRIT*.

TLS *ESPRIT* SOLUTION

The TLS problem associated with *ESPRIT* is to find $\mathbf{\Psi}$ in the equation

$$(\mathbf{B} - \mathbf{\Delta})\mathbf{\Psi} = \mathbf{B}' - \mathbf{\Delta}'$$

to minimize $\left\| \begin{array}{cc} \mathbf{\Delta} & \mathbf{\Delta}' \end{array} \right\|_F$.

(This is a generalization of the TLS problem described earlier.)

Define $\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$

as the matrix of *right singular vectors* of the matrix $\begin{bmatrix} \mathbf{B} & \mathbf{B}' \end{bmatrix}$.

The TLS solution is given by $\mathbf{\Psi}_{TLS} = -\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$.

ESPRIT ALGORITHM SUMMARY

1. Define the $N+1$-dimensional random vector $\bar{\boldsymbol{x}}$ pertaining to $N+1$ consecutive data samples $x[0], x[1], \ldots, x[N]$ and estimate the correlation matrix $\hat{\mathbf{R}}_{\bar{\boldsymbol{x}}}$ from the data. [Usually the covariance method or the modified covariance method should be used here, especially if the total length of the data record (N_s) is small.]

2. Compute the generalized eigenvectors and eigenvalues of $\hat{\mathbf{R}}_{\bar{\boldsymbol{x}}}$:

$$\hat{\mathbf{R}}_{\bar{\boldsymbol{x}}}\bar{\mathbf{e}}_k = \bar{\lambda}_k \mathbf{\Sigma}_{\bar{\boldsymbol{\eta}}}\bar{\mathbf{e}}_k \qquad\qquad k = 1, 2, \ldots, N+1$$

3. If necessary, estimate the number of signals M.

ESPRIT (cont'd.)

4. Generate a basis spanning the signal subspace and partition it as

$$\bar{\mathbf{B}} = \mathbf{\Sigma}_{\bar{\boldsymbol{\eta}}} \begin{bmatrix} | & & | \\ \bar{\mathbf{e}}_1 & \cdots & \bar{\mathbf{e}}_M \\ | & & | \end{bmatrix} = \begin{bmatrix} \mathbf{B} \\ \\ \times \cdots \times \end{bmatrix} = \begin{bmatrix} \times \cdots \times \\ \\ \mathbf{B}' \end{bmatrix}$$

5. Compute the matrix $\mathbf{V}$ of right singular vectors of

$$\begin{bmatrix} \mathbf{B} & \mathbf{B}' \end{bmatrix}$$

and partition $\mathbf{V}$ into four $M \times M$ submatrices

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix}$$

ESPRIT (cont'd.)

6. Compute the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_M$ of the matrix $\mathbf{\Psi}_{TLS} = -\mathbf{V}_{12}\mathbf{V}_{22}^{-1}$.

7. Find the desired frequencies as

$$\omega_k = \angle\lambda_k \qquad\qquad k = 1, 2, \ldots, M$$

"SIGNAL COPY" FEATURE OF *ESPRIT*

- Estimates for the signal amplitudes can be made using

$$\begin{bmatrix} \widehat{A}_1 \\ \widehat{A}_2 \\ \vdots \\ \widehat{A}_M \end{bmatrix} = \mathbf{W}_{SC}^{*T} \bar{\boldsymbol{x}}$$

where

$$\mathbf{W}_{SC} = \bar{\mathbf{S}} \left(\bar{\mathbf{S}}^{*T} \bar{\mathbf{S}} \right)^{*T}$$

- Estimates approach true values as $\sigma_o^2 \rightarrow 0$.

COMPUTATIONAL CONSIDERATIONS FOR SUBSPACE METHODS

- Avoiding computation of the correlation matrix
  - Use of the data matrix and SVD

- Statistical methods for estimating the number of signals

- Evaluating the pseudospectrum

AVOIDING THE CORRELATION MATRIX

- Eigenvalues/vectors can be found from the SVD of the data matrix $\mathbf{X}$ (assuming $\hat{\mathbf{R}}_{\boldsymbol{x}} = \mathbf{X}^{*T}\mathbf{X}$)

- For PCLP the filter coefficients can be found from

$$\mathbf{a}' = -\mathbf{X}_1^{+(M)}\mathbf{x}_0 \quad \text{where} \quad \mathbf{X} = \begin{bmatrix} | & \\ \mathbf{x}_0 & \mathbf{X}_1 \\ | & \end{bmatrix}$$

and $\mathbf{X}_1^{+(M)}$ denotes the rank M pseudoinverse of $\mathbf{X}_1$.

For the Kumaresan-Prony case a simple computation is

$$\mathbf{X}_1^{+(M)} = \mathbf{X}_1^{+} = \mathbf{X}_1^{*T}(\mathbf{X}_1\mathbf{X}_1^{*T})^{-1}$$

STATISTICAL ESTIMATION OF M

1. Find M to minimize either AIC or MDL:

$$\begin{aligned} \mathsf{AIC}(M) &= -2K(N-M)\ln\varrho(M) + 2M(2N-M) \\ \mathsf{MDL}(M) &= -K(N-M)\ln\varrho(M) + \tfrac{1}{2}M(2N-M)\ln K \end{aligned}$$

   where

$$\varrho(M) = \frac{(\lambda_{M+1}\lambda_{M+2}\cdots\lambda_N)^{\frac{1}{N-M}}}{\frac{1}{N-M}\left(\lambda_{M+1}+\lambda_{M+2}+\cdots+\lambda_N\right)}$$

2. Estimate the noise power as

$$\hat{\sigma}_o^2 = \frac{1}{N-M}\left(\lambda_{M+1}+\lambda_{M+2}+\cdots+\lambda_N\right)$$

COMPUTING THE PSEUDOSPECTRUM

- For estimates involving a single vector, use FFT of vector.

  Example:

  $$\widehat{P}_{MN}(e^{\jmath\omega}) = \frac{1}{|\mathbf{w}^{*T}\mathbf{d}|^2} \qquad \text{Compute FFT}\{d[n]\}$$

- For estimates involving a matrix, such as

  $$\widehat{P}_{MU}(e^{\jmath\omega}) = \frac{1}{\mathbf{w}^{*T}\boldsymbol{P}_{noise}\mathbf{w}}$$

  use procedure similar to computation of the ML spectral estimate.